



INTERNATIONAL APPLICATION PUBLISHED UNDER THE PATENT COOPERATION TREATY (PCT)

(51) International Patent Classification ⁶: C12N 15/12, C07K 14/47, C12N 16/70, 1/21, A61K 38/17, C07K 16/18, C12Q 1/68	A2	(11) International Publication Number: WO 99/15658 (43) International Publication Date: 1 April 1999 (01.04.99)
(21) International Application Number: PCT/US98/19839 (22) International Filing Date: 22 September 1998 (22.09.98) (30) Priority Data: 08/933,750 23 September 1997 (23.09.97) US (63) Related by Continuation (CON) or Continuation-in-Part (CIP) to Earlier Application US 08/933,750 (CIP) Filed on 23 September 1997 (23.09.97) (71) Applicant (for all designated States except US): INCYTE PHARMACEUTICALS, INC. [US/US]; 3174 Porter Drive, Palo Alto, CA 94304 (US). (72) Inventors; and (75) Inventors/Applicants (for US only): LAL, Preeti [IN/US]; 2382 Lass Drive, Santa Clara, CA 95054 (US). HILLMAN, Jennifer, L. [US/US]; 230 Monroe Drive #12, Mountain View, CA 94040 (US). BANDMAN, Olga [US/US]; 366 Anna Avenue, Mountain View, CA 94043 (US). SHAH, Purvi [IN/US]; 1608 Queen Charlotte Drive #5, Sunnyvale, CA 94087 (US). AU-YOUNG, Janice [US/US]; 1419 Kains		Avenue, Berkeley, CA 94702 (US). YUE, Henry [US/US]; 826 Lois Avenue, Sunnyvale, CA 94087 (US). GUEGLER, Karl, J. [CH/US]; 1048 Oakland Avenue, Menlo Park, CA 94025 (US). (74) Agents: BILLINGS, Lucy, J. et al.; Incyte Pharmaceuticals, Inc., 3174 Porter Drive, Palo Alto, CA 94304 (US). (81) Designated States: AL, AM, AT, AU, AZ, BA, BB, BG, BR, BY, CA, CH, CN, CU, CZ, DE, DK, EE, ES, FI, GB, GE, GH, GM, HR, HU, ID, IL, IS, JP, KE, KG, KP, KR, KZ, LC, LK, LR, LS, LT, LU, LV, MD, MG, MK, MN, MW, MX, NO, NZ, PL, PT, RO, RU, SD, SE, SG, SI, SK, SL, TJ, TM, TR, TT, UA, UG, US, UZ, VN, YU, ZW, ARIPO patent (GH, GM, KE, LS, MW, SD, SZ, UG, ZW), Eurasian patent (AM, AZ, BY, KG, KZ, MD, RU, TJ, TM), European patent (AT, BE, CH, CY, DE, DK, ES, FI, FR, GB, GR, IE, IT, LU, MC, NL, PT, SE), OAPI patent (BF, BJ, CF, CG, CI, CM, GA, GN, GW, ML, MR, NE, SN, TD, TG). Published <i>Without international search report and to be republished upon receipt of that report.</i>
(54) Title: HUMAN REGULATORY MOLECULES (57) Abstract The invention provides human regulatory molecules and polynucleotides (collectively designated HRM) which identify and encode them. The invention also provides expression vectors, host cells, agonists, antibodies and antagonists. The invention further provides methods for diagnosing, preventing, and treating disorders associated with expression of human regulatory molecules.		

FOR THE PURPOSES OF INFORMATION ONLY

Codes used to identify States party to the PCT on the front pages of pamphlets publishing international applications under the PCT.

AL	Albania	ES	Spain	LS	Lesotho	SI	Slovenia
AM	Armenia	FI	Finland	LT	Lithuania	SK	Slovakia
AT	Austria	FR	France	LU	Luxembourg	SN	Senegal
AU	Australia	GA	Gabon	LV	Latvia	SZ	Swaziland
AZ	Azerbaijan	GB	United Kingdom	MC	Monaco	TD	Chad
BA	Bosnia and Herzegovina	GE	Georgia	MD	Republic of Moldova	TG	Togo
BB	Barbados	GH	Ghana	MG	Madagascar	TJ	Tajikistan
BE	Belgium	GN	Guinea	MK	The former Yugoslav Republic of Macedonia	TM	Turkmenistan
BF	Burkina Faso	GR	Greece	ML	Mali	TR	Turkey
BG	Bulgaria	HU	Hungary	MN	Mongolia	TT	Trinidad and Tobago
BJ	Benin	IE	Ireland	MR	Mauritania	UA	Ukraine
BR	Brazil	IL	Israel	MW	Malawi	UG	Uganda
BY	Belarus	IS	Iceland	MX	Mexico	US	United States of America
CA	Canada	IT	Italy	NE	Niger	UZ	Uzbekistan
CF	Central African Republic	JP	Japan	NL	Netherlands	VN	Viet Nam
CG	Congo	KE	Kenya	NO	Norway	YU	Yugoslavia
CH	Switzerland	KG	Kyrgyzstan	NZ	New Zealand	ZW	Zimbabwe
CI	Côte d'Ivoire	KP	Democratic People's Republic of Korea	PL	Poland		
CM	Cameroon	KR	Republic of Korea	PT	Portugal		
CN	China	KZ	Kazakstan	RO	Romania		
CU	Cuba	LC	Saint Lucia	RU	Russian Federation		
CZ	Czech Republic	LI	Liechtenstein	SD	Sudan		
DE	Germany	LK	Sri Lanka	SE	Sweden		
DK	Denmark	LR	Liberia	SG	Singapore		
EE	Estonia						

HUMAN REGULATORY MOLECULES

5

TECHNICAL FIELD

This invention relates to nucleic acid and amino acid sequences of human regulatory molecules which are important in disease and to the use of these sequences in the diagnosis, prevention, and treatment of diseases associated with cell proliferation.

10

BACKGROUND OF THE INVENTION

Cells grow and differentiate, carry out their structural or metabolic roles, participate in organismal development, and respond to their environment by altering their gene expression. Cellular functions are controlled by the timing and amount of expression attributable to thousands of individual genes. The regulation of expression is metabolically vital in that it
15 conserves energy and prevents the synthesis and accumulation of intermediates such as RNA and incomplete or inactive proteins when the gene product is not needed.

Regulatory protein molecules are absolutely essential in the control of gene expression. These molecules turn individual or groups of genes on and off in response to various inductive mechanisms of the cell or organism; act as transcription factors by
20 determining whether or not transcription is initiated, enhanced, or repressed; and splice transcripts as dictated in a particular cell or tissue. Although regulatory molecules interact with short stretches of DNA scattered throughout the entire genome, most gene expression is regulated near the site at which transcription starts or within the open reading frame of the gene being expressed. The regulated stretches of the DNA can be simple and interact with
25 only a single protein, or they can require several proteins acting as part of a complex to order to regulate gene expression.

The double helix structure and repeated sequences of DNA create external features which can be recognized by the regulatory molecules. These external features are hydrogen bond donor and acceptor groups, hydrophobic patches, major and minor grooves, and regular,
30 repeated stretches of sequence which cause distinct bends in the helix. Such features provide recognition sites for the binding of regulatory proteins. Typically, these recognition sites are less than 20 nucleotides in length although multiple sites may be adjacent to each other and

each may exert control over a single gene. Hundreds of these DNA sequences have been identified, and each is recognized by a different protein or complex of proteins which carry out gene regulation.

The regulatory protein molecules or complexes recognize and bind to specific
5 nucleotide sequences of upstream (5') nontranslated regions, which precede the first translated exon of the open reading frame (ORF); of intron junctions, which occur between the many exons of the OR; and of downstream (3') untranslated regions, which follow the ORF. The regulatory molecule surface features are extensively complementary to the surface features of the double helix. Even though each individual contact between the protein(s) and
10 helix may be relatively weak (hydrogen bonds, ionic bonds, and/or hydrophobic interactions) and the 20 or more contacts occurring between the protein and DNA result in a highly specific and very strong interaction.

Families of regulatory molecules

Many of the regulatory molecules incorporate one of a set of DNA-binding structural
15 motifs, each of which contains either α helices or β sheets and binds to the major groove of DNA. Seven of the structural motifs common to regulatory molecules are helix-turn-helix, homeodomains, zinc finger, steroid receptor, β sheets, leucine zipper, and helix-loop-helix.

The helix-turn-helix motif is constructed from two α helices connected by a short chain of amino acids, which constitutes the "turn". The two helices interact with each other
20 to form a fixed angle. The more carboxy-terminal helix is called the recognition helix because it fits into the major groove of the DNA. The amino acid side chains of the helix recognize the specific DNA sequence to which the protein binds. The remaining structure varies a great deal among the regulatory proteins incorporating this motif. The helix-turn-helix configuration is not stable without the rest of the protein and will not bind to DNA
25 without other peptide regions providing stability. Other peptide regions also interact with the DNA, increasing the number of unique sequences a helix-turn-helix can recognize.

Many sequence-specific DNA binding proteins actually bind as symmetric dimers to DNA sequences that are composed of two very similar half-sites, also arranged symmetrically. This configuration allows each protein monomer to interact in the same way
30 with the DNA recognition site and doubles the number of contacts with the DNA. This doubling of contacts greatly increases the binding affinity while only doubling the free energy of the interaction. Helix-turn-helix motifs always bind to DNA that is in the B-DNA form.

The homeodomain motif is found in a special group of helix-turn-helix proteins that are encoded by homeotic selector genes, so called because the proteins encoded by these genes control developmental switches. For example, mutations in these genes cause one body part to be converted into another in the fruit fly, Drosophila. These genes have been found in every eukaryotic organism studied. The helix-turn-helix region of different homeodomains is always surrounded by the same structure, but not necessarily the same sequence, and the motif is always presented to DNA the same way. This helix-turn-helix configuration is stable by itself and, when isolated, can still bind to DNA. It may be significant that the helices in homeodomains are generally longer than the helices in most HLH regulatory proteins.

Portions of the motif which interact most directly with DNA differ among these two families. Detailed examples of DNA-protein binding are described in Pabo, C.O. and R.T. Sauer (1992; Ann. Rev. Biochem. 61:1053-95).

A third motif incorporates zinc molecules into the crucial portion of the protein. These proteins are most often referred to as having zinc fingers, although their structure can be one of several types. Proteins in this family often contain tandem repeats of the 30-residue zinc finger motif, including the sequence patterns Cys-X2 or 4-Cys-X12-His-X3-5-His. Each of these regulatory proteins has an α helix and an antiparallel β sheet. Two histidines in the α helix and 2 cysteines near the turn in the β sheet interact with the zinc ion which holds the α helix and the β sheet together. Contact with the DNA is made by the arginine preceding the α helix, and by the second, third, and sixth residues of the α helix. When this arrangement is repeated as a cluster of several fingers, the α helix of each finger can contact and interact with the major groove of the DNA. By changing the number of zinc fingers, the specificity and strength of the binding interaction can be altered.

The steroid receptors are a family of intracellular proteins that include receptors for steroids, retinoids, vitamin D, thyroid hormones, and other important compounds. The DNA binding domain of these proteins contains about 70 residues, eight of which are conserved cysteines. The steroid receptor motif forms a structure in which two α helices are packed perpendicularly to each other, forming more of a globular shape than a finger. Each helix has a zinc ion which holds a peptide loop against the N-terminal end of the helix. The first helix fits into the major groove of DNA, and side chains make contacts with edges of the DNA base pairs. The steroid receptor proteins, like the helix-turn-helix proteins, form dimers that bind the DNA. The second helix of each monomer contacts the phosphate groups of the

DNA backbone and also provides the dimerization interface. In some cases, multiple choices can exist for heterodimerization which produces another mechanism for fine-tuning the regulation of numerous genes.

Another family of regulatory protein molecules uses a motif consisting of a two-
5 stranded antiparallel β sheet to recognize the major groove of DNA. The exact DNA sequence recognized by the motif depends on the amino acid sequence in the β sheet from which the amino acid side chains extend and contact the DNA. In two prokaryotic examples of the β sheet, the regulatory proteins form tetramers when binding DNA.

The leucine zipper motif commonly forms dimers and has a 30-40 residue motif in
10 which two α helices (one from each monomer) are joined to form a short coiled-coil. The helices are held together by interactions among hydrophobic amino acid side chains (often on heptad-repeated leucines) that extend from one side of each helix. Beyond this, the helices separate, and each basic region contacts the major groove of DNA. Proteins with the leucine zipper motif can also form either homodimers or heterodimers, thus extending the specific
15 combinations available to activate or repress expression.

Yet another important motif is the helix-loop-helix, which consists of a short α helix connected by a loop to a longer α helix. The loop is flexible and allows the two helices to fold back against each other. The α helices bind both to DNA and to the HLH structure of another protein. The second protein can be the same (producing homodimers) or different
20 (producing heterodimers). Some HLH monomers lack sufficient α helix to bind DNA, but they can still form heterodimers which can serve to inactivate specific regulatory proteins.

Hundreds of regulatory proteins have been identified to date, and more are being characterized in a wide variety of organisms. Most regulatory proteins have at least one of the common structural motifs for making contact with DNA, but several important regulatory
25 proteins, such as the p53 tumor suppressor gene, do not share their structure with other known regulatory proteins. Variations on known motifs and new motifs have been and are currently being characterized (Faisst, S. and S. Meyer (1992) Nucl. Acids Res. 20: 3-26).

Although binding of DNA to a regulatory protein is very specific, there is no way to predict the exact DNA sequence to which a particular regulatory protein will bind or the
30 primary structure of a regulatory protein for a specific DNA sequence. Thus, interactions of DNA and regulatory proteins are not limited to the motifs described above. Other domains of the proteins often form crucial contacts with the DNA, and accessory proteins can provide

important interactions which may convert a particular protein complex to an activator, or a repressor, or may prevent binding (Alberts, B. et al. (1994) Molecular Biology of the Cell, Garland Publishing Co, New York NY, pp. 401-74).

Diseases and disorders related to gene regulation

5 Many neoplastic growths in humans can be traced to problems of gene regulation. Malignant growth of cells may be the result of excess transcriptional activator or loss of an inhibitor or suppressor (Cleary ML (1992) Cancer Surv. 15:89-104). Alternatively, gene fusion may produce chimeric loci with switched domains, such that the level of activation is no longer correct for the gene specificity of that factor.

10 The cellular response to infection or trauma is beneficial when gene expression is appropriate. However, when hyper-responsivity or another imbalance occurs for any reason, improper or insufficient regulation of gene expression may cause considerable tissue or organ damage. This damage is well documented in immunological responses to allergens, heart attack, stroke, and infections (Harrison's Principles of Internal Medicine, 13/e^o, (1994)
15 McGraw Hill, Inc. and Teton Data Systems Software). In addition, the accumulation of somatic mutations and the increasing inability to regulate cellular responses is seen in the prevalence of osteoarthritis and onset of other disorders associated with aging.

The discovery of new human regulatory protein molecules which are important in disease development and the polynucleotides encoding them satisfies a need in the art by
20 providing new compositions which are useful in the diagnosis, prevention and treatment of diseases associated with cell proliferation, particularly immune responses and cancers.

SUMMARY OF THE INVENTION

The invention features substantially purified polypeptides, human regulatory molecules, collectively referred to as HRM and individually referred to as HRM-1 through
25 HRM-10, having at least one of the amino acid sequences selected from the group consisting of SEQ ID NOs:1-10.

The invention further provides isolated and substantially purified polynucleotide sequences encoding HRM. In a particular aspect, the polynucleotide is at least one the nucleotide sequences selected from the group consisting of SEQ ID NOs:11-20.

30 In addition, the invention provides a polynucleotide sequence, or fragment thereof, which hybridizes under stringent conditions to any of the polynucleotide sequences of SEQ ID NOs:11-20. In another aspect the invention provides compositions comprising isolated

and purified polynucleotide sequences of SEQ ID NOs: 11-20 or fragments thereof.

The invention further provides a polynucleotide sequence comprising the complement or fragments thereof of any one of the polynucleotide sequences encoding HRM. In another aspect the invention provides compositions comprising isolated and purified polynucleotide sequences comprising the complements of SEQ ID NOs: 11-20 or fragments thereof.

The present invention further provides an expression vector containing at least a fragment of any one of the polynucleotide sequences of SEQ ID NOs: 11-20. In yet another aspect, the expression vector containing the polynucleotide sequence is contained within a host cell.

The invention also provides a method for producing a polypeptide or a fragment thereof, the method comprising the steps of: a) culturing the host cell containing an expression vector containing at least a fragment of the polynucleotide sequence encoding an HRM under conditions suitable for the expression of the polypeptide; and b) recovering the polypeptide from the host cell culture.

The invention also provides a pharmaceutical composition comprising a substantially purified HRM in conjunction with a suitable pharmaceutical carrier.

The invention also provides a purified antagonist of HRM. In one aspect the invention provides a purified antibody which binds to an HRM.

Still further, the invention provides a purified agonist of HRM.

The invention also provides a method for treating or preventing a cancer comprising administering to a subject in need of such treatment an effective amount of a pharmaceutical composition containing HRM.

The invention also provides a method for treating or preventing a cancer comprising administering to a subject in need of such treatment an effective amount of an antagonist of HRM.

The invention also provides a method for treating or preventing an immune response associated with the increased expression or activity of HRM comprising administering to a subject in need of such treatment an effective amount of an antagonist of HRM.

The invention also provides a method for stimulating cell proliferation comprising administering to a cell an effective amount of purified HRM.

The invention also provides a method for detecting a polynucleotide which encodes a human regulatory molecule in a biological sample comprising the steps of: a) hybridizing a

polynucleotide sequence complementary to a polynucleotide encoding HRM to nucleic acid material of a biological sample, thereby forming a hybridization complex; and b) detecting the hybridization complex, wherein the presence of the complex correlates with the presence of a polynucleotide encoding the human regulatory molecule in the biological sample.

5 The invention also provides a microarray which contains at least a fragment of at least one of the polynucleotide sequences encoding HRM. In a particular aspect, the microarray contains at least a fragment of at least one of the sequences selected from the group consisting of SEQ ID NOs:11-20.

10 The invention also provides a method for the simultaneous detection of the levels of expression of polynucleotides which encode human regulatory molecules in a biological sample comprising the steps of: a) hybridizing said microarray to labeled complementary nucleotides of a biological sample, comprising at least a fragment of at least one of the polynucleotides encoding HRM, thereby forming hybridization complexes; and b) quantifying expression, wherein the signal produced by the hybridization complexes
15 correlates with expression of particular polynucleotides encoding human regulatory molecules in the biological sample. In a preferred embodiment, prior to hybridization, the nucleic acid material of the biological sample is amplified and labeled by the polymerase chain reaction.

DESCRIPTION OF THE INVENTION

20 Before the present proteins, nucleotide sequences, and methods are described, it is understood that this invention is not limited to the particular methodology, protocols, cell lines, vectors, and reagents described, as these may vary. It is also to be understood that the terminology used herein is for the purpose of describing particular embodiments only, and is not intended to limit the scope of the present invention which will be limited only by the appended claims.

25 It must be noted that as used herein and in the appended claims, the singular forms "a", "an", and "the" include plural reference unless the context clearly dictates otherwise. Thus, for example, reference to "a host cell" includes a plurality of such host cells, reference to the "antibody" is a reference to one or more antibodies and equivalents thereof known to those skilled in the art, and so forth.

30 Unless defined otherwise, all technical and scientific terms used herein have the same meanings commonly understood by one of ordinary skill in the art to which this invention belongs. Although any methods and materials similar or equivalent to those described herein

can be used in the practice or testing of the present invention, the preferred methods, devices, and materials are now described. All publications mentioned herein are incorporated herein by reference for the purpose of describing and disclosing the cell lines, vectors, arrays and methodologies which are reported in the publications which might be used in connection with the invention. Nothing herein is to be construed as an admission that the invention is not entitled to antedate such disclosure by virtue of prior invention.

DEFINITIONS

HRM refers to the amino acid sequences of substantially purified HRM obtained from any species, particularly mammalian, including bovine, ovine, porcine, murine, equine, and preferably human, from any source whether natural, synthetic, semi-synthetic, or recombinant.

"Agonist" refers to a molecule which, when bound to HRM, increases or prolongs the duration of the effect of HRM. Agonists may include proteins, nucleic acids, carbohydrates, or any other molecules which bind to and modulate the effect of HRM.

An "allele" or "allelic sequence" is an alternative form of the gene encoding HRM. Alleles may result from at least one mutation in the nucleic acid sequence and may result in altered mRNAs or polypeptides whose structure or function may or may not be altered. Any given natural or recombinant gene may have none, one, or many allelic forms. Common mutational changes which give rise to alleles are generally ascribed to natural deletions, additions, or substitutions of nucleotides. Each of these types of changes may occur alone, or in combination with the others, one or more times in a given sequence.

"Altered" nucleic acid sequences encoding HRM include those with deletions, insertions, or substitutions of different nucleotides resulting in a polynucleotide that encodes the same or a functionally equivalent HRM. Included within this definition are polymorphisms which may or may not be readily detectable using a particular oligonucleotide probe of the polynucleotide encoding HRM, and improper or unexpected hybridization to alleles, with a locus other than the normal chromosomal locus for the polynucleotide sequence encoding HRM. The encoded protein may also be "altered" and contain deletions, insertions, or substitutions of amino acid residues which produce a silent change and result in a functionally equivalent HRM. Deliberate amino acid substitutions may be made on the basis of similarity in polarity, charge, solubility, hydrophobicity, hydrophilicity, and/or the amphipathic nature of the residues as long as the biological or immunological activity of

HRM is retained. For example, negatively charged amino acids may include aspartic acid and glutamic acid; positively charged amino acids may include lysine and arginine; and amino acids with uncharged polar head groups having similar hydrophilicity values may include leucine, isoleucine, and valine, glycine and alanine, asparagine and glutamine, serine and threonine, and phenylalanine and tyrosine.

"Amino acid sequence" refers to an oligopeptide, peptide, polypeptide, or protein sequence, and fragment thereof, and to naturally occurring or synthetic molecules. Fragments of HRM are preferably about 5 to about 15 amino acids in length and retain the biological activity or the immunological activity of HRM. Where "amino acid sequence" is recited herein to refer to an amino acid sequence of a naturally occurring protein molecule, amino acid sequence, and like terms, are not meant to limit the amino acid sequence to the complete, native amino acid sequence associated with the recited protein molecule.

"Amplification" refers to the production of additional copies of a nucleic acid sequence and is generally carried out using polymerase chain reaction (PCR) technologies well known in the art

"Antagonist" refers to a molecule which, when bound to HRM, decreases the amount or the duration of the effect of the biological or immunological activity of HRM. Antagonists may include proteins, nucleic acids, carbohydrates, or any other molecules which decrease the effect of HRM.

"Antibody" refers to intact molecules as well as fragments thereof, such as F_a , $F(ab')_2$, and F_v , which are capable of binding the epitopic determinant. Antibodies that bind HRM polypeptides can be prepared using intact polypeptides or fragments containing small peptides of interest as the immunizing antigen. The polypeptide or oligopeptide used to immunize an animal can be derived from the translation of RNA or synthesized chemically and can be conjugated to a carrier protein, if desired. Commonly used carriers that are chemically coupled to peptides include bovine serum albumin and thyroglobulin, keyhole limpet hemocyanin. The coupled peptide is then used to immunize the animal (e.g., a mouse, a rat, or a rabbit).

"Antigenic determinant" refers to that fragment of a molecule (i.e., an epitope) that makes contact with a particular antibody. When a protein or fragment of a protein is used to immunize a host animal, numerous regions of the protein may induce the production of antibodies which bind specifically to a given region or three-dimensional structure on the

protein; these regions or structures are referred to as antigenic determinants. An antigenic determinant may compete with the intact antigen (i.e., the immunogen used to elicit the immune response) for binding to an antibody.

“Antisense” refers to any composition containing nucleotide sequences which are complementary to a specific DNA or RNA sequence. The term “antisense strand” is used in reference to a nucleic acid strand that is complementary to the “sense” strand. Antisense molecules include peptide nucleic acids and may be produced by any method including synthesis or transcription. Once introduced into a cell, the complementary nucleotides combine with natural sequences produced by the cell to form duplexes and block either transcription or translation. The designation “negative” is sometimes used in reference to the antisense strand, and “positive” is sometimes used in reference to the sense strand.

“Biologically active” refers to a protein having structural, regulatory, or biochemical functions of a naturally occurring molecule. Likewise, “immunologically active” refers to the capability of the natural, recombinant, or synthetic HRM, or any oligopeptide thereof, to induce a specific immune response in appropriate animals or cells and to bind with specific antibodies.

“Complementary” or “complementarity” refer to the natural binding of polynucleotides under permissive salt and temperature conditions by base-pairing. For example, the sequence “A-G-T” binds to the complementary sequence “T-C-A”. Complementarity between two single-stranded molecules may be “partial”, in which only some of the nucleic acids bind, or it may be complete when total complementarity exists between the single stranded molecules. The degree of complementarity between nucleic acid strands has significant effects on the efficiency and strength of hybridization between nucleic acid strands. This is of particular importance in amplification reactions, which depend upon binding between nucleic acids strands and in the design and use of PNA molecules.

A “composition comprising a given polynucleotide sequence” refers broadly to any composition containing the given polynucleotide sequence. The composition may comprise a dry formulation or an aqueous solution. Compositions comprising polynucleotide sequences encoding HRM (SEQ ID NOs: 11-20) or fragments thereof may be employed as hybridization probes. The probes may be stored in freeze-dried form and may be associated with a stabilizing agent such as a carbohydrate. In hybridizations, the probe may be deployed in an aqueous solution containing salts (e.g., NaCl), detergents (e.g., SDS) and other components

(e.g., Denhardt's solution, dry milk, salmon sperm DNA, etc.).

"Consensus" refers to a nucleic acid sequence which has been resequenced to resolve uncalled bases, has been extended using XL-PCR kit (Perkin Elmer, Norwalk CT) in the 5' and/or the 3' direction and resequenced, or has been assembled from the overlapping
5 sequences of more than one Incyte Clone using a computer program for fragment assembly (e.g., GELVIEW Fragment Assembly system, GCG, Madison WI). Some sequences have been both extended and assembled to produce the consensus sequence .

The phrase "correlates with expression of a polynucleotide" indicates that the detection of the presence of a ribonucleic acid that is similar to a polynucleotide encoding an
10 HRM by northern analysis is indicative of the presence of mRNA encoding HRM in a sample and thereby correlates with expression of the transcript from the polynucleotide encoding the protein.

"HRM" refers to any or all of the human polypeptides, HRM-1 through HRM-10.

"Deletion" refers to a change in the amino acid or nucleotide sequence and results in
15 the absence of one or more amino acid residues or nucleotides.

"Derivative" refers to the chemical modification of a nucleic acid encoding or complementary to HRM or the encoded HRM. Such modifications include, for example, replacement of hydrogen by an alkyl, acyl, or amino group. A nucleic acid derivative encodes a polypeptide which retains the biological or immunological function of the natural molecule.
20 A derivative polypeptide is one which is modified by glycosylation, pegylation, or any similar process which retains the biological or immunological function of the polypeptide from which it was derived.

"Homology" refers to a degree of complementarity. There may be partial homology or complete homology (i.e., identity). A partially complementary sequence that at least
25 partially inhibits an identical sequence from hybridizing to a target nucleic acid is referred to using the functional term "substantially homologous." The inhibition of hybridization of the completely complementary sequence to the target sequence may be examined using a hybridization assay (Southern or northern blot, solution hybridization and the like) under conditions of low stringency. A substantially homologous sequence or hybridization probe
30 will compete for and inhibit the binding of a completely homologous sequence to the target sequence under conditions of low stringency. This is not to say that conditions of low stringency are such that non-specific binding is permitted; low stringency conditions require

that the binding of two sequences to one another be a specific (i.e., selective) interaction. The absence of non-specific binding may be tested by the use of a second target sequence which lacks even a partial degree of complementarity (e.g., less than about 30% identity). In the absence of non-specific binding, the probe will not hybridize to the second non-complementary target sequence.

Human artificial chromosomes (HACs) are linear microchromosomes which may contain DNA sequences of 10K to 10M in size and contain all of the elements required for stable mitotic chromosome segregation and maintenance.

"Humanized antibody" refers to antibody molecules in which amino acids have been replaced in the non-antigen binding regions in order to more closely resemble a human antibody, while still retaining the original binding ability.

"Hybridization" refers to any process by which a strand of nucleic acid bonds with a complementary strand through base pairing.

"Hybridization complex" refers to a complex formed between two nucleic acid sequences by virtue of the formation of hydrogen bonds between complementary G and C bases and between complementary A and T bases; these hydrogen bonds may be further stabilized by base stacking interactions. The two complementary nucleic acid sequences hydrogen bond in an antiparallel configuration. A hybridization complex may be formed in solution (e.g., C_0t or R_0t analysis) or between one nucleic acid sequence present in solution and another nucleic acid sequence immobilized on a solid support (e.g., paper, membranes, filters, chips, pins or glass slides, or any other appropriate substrate to which cells or their nucleic acids have been fixed).

An "insertion" or "addition" refers to a change in an amino acid or nucleotide sequence resulting in the addition of one or more amino acid residues or nucleotides, respectively, as compared to the naturally occurring molecule.

"Microarray" refers to an array (or arrangement) of distinct oligonucleotides synthesized on a substrate, such as paper, nylon or other type of membrane, filter, gel, polymer, chip, glass slide, or any other suitable support.

"Modulate" refers to a change in the activity of HRM. For example, modulation may cause an increase or a decrease in protein activity, binding characteristics, or any other biological, functional or immunological properties of HRM.

"Nucleic acid sequence" refers to an oligonucleotide, nucleotide, or polynucleotide,

and fragments thereof, and to DNA or RNA of genomic or synthetic origin which may be single- or double-stranded, and represent the sense or antisense strand. "Fragments" are those nucleic acid sequences which are greater than 60 nucleotides in length, and most preferably includes fragments that are at least 100 nucleotides or at least 1000 nucleotides, and at least 10,000 nucleotides in length.

"Oligonucleotide" refers to a nucleic acid sequence of at least about 6 nucleotides to about 60 nucleotides, preferably about 15 to 30 nucleotides, and more preferably about 20 to 25 nucleotides, which can be used in PCR amplification or hybridization assays. As used herein, oligonucleotide is substantially equivalent to the terms "amplimers", "primers", "oligomers", and "probes", as commonly defined in the art.

"Peptide nucleic acid" (PNA) refers to an antisense molecule or anti-gene agent which comprises an oligonucleotide of at least five nucleotides in length linked to a peptide backbone of amino acid residues which ends in lysine. The terminal lysine confers solubility to the composition. PNAs may be pegylated to extend their lifespan in the cell where they preferentially bind complementary single stranded DNA and RNA and stop transcript elongation (Nielsen, P.E. et al. (1993) Anticancer Drug Des. 8:53-63).

"Portion", with regard to a protein (as in "a portion of a given protein"), refers to fragments of that protein. The fragments may range in size from five amino acid residues to the entire amino acid sequence minus one amino acid. Thus, a protein "comprising at least a portion of the amino acid sequence of an HRM encompasses the full-length HRM and fragments thereof.

"Sample" is used in its broadest sense. A biological sample suspected of containing nucleic acid encoding HRM, or fragments thereof, or HRM itself may comprise a bodily fluid, extract from a cell, chromosome, organelle, or membrane isolated from a cell, a cell, genomic DNA, RNA, or cDNA (in solution or bound to a solid support, a tissue, a tissue print, and the like.

"Specific binding" or "specifically binding" refers to that interaction between a protein or peptide and an agonist, an antibody and an antagonist. The interaction is dependent upon the presence of a particular structure (i.e., the antigenic determinant or epitope) of the protein recognized by the binding molecule. For example, if an antibody is specific for epitope "A", the presence of a protein containing epitope A (or free, unlabeled A) in a reaction containing labeled "A" and the antibody will reduce the amount of labeled A bound

to the antibody.

"Stringent conditions" or "stringency" refer to the conditions for hybridization as defined by the nucleic acid, salt, and temperature. These conditions are well known in the art and may be altered in order to identify or detect identical or related polynucleotide sequences.

5 Numerous equivalent conditions comprising either low or high stringency depend on factors such as the length and nature of the sequence (DNA, RNA, base composition), nature of the target (DNA, RNA, base composition), milieu (in solution or immobilized on a solid substrate), concentration of salts and other components (e.g., formamide, dextran sulfate and/or polyethylene glycol), and temperature of the reactions (within a range from about 5°C
10 below the melting temperature of the probe to about 20°C to 25°C below the melting temperature). One or more factors may be varied to generate conditions of either low or high stringency different from, but equivalent to, the above listed conditions.

"Substantially purified" refers to nucleic or amino acid sequences that are removed from their natural environment, isolated or separated, and are at least 60% free, preferably
15 75% free, and most preferably 90% free from other components with which they are naturally associated.

A "substitution" refers to the replacement of one or more amino acids or nucleotides by different amino acids or nucleotides, respectively.

"Transformation" describes a process by which exogenous DNA enters and changes a
20 recipient cell. It may occur under natural or artificial conditions using various methods well known in the art. Transformation may rely on any known method for the insertion of foreign nucleic acid sequences into a prokaryotic or eukaryotic host cell. The method is selected based on the type of host cell being transformed and may include, but is not limited to, viral infection, electroporation, heat shock, lipofection, and particle bombardment. Such
25 "transformed" cells include stably transformed cells in which the inserted DNA is capable of replication either as an autonomously replicating plasmid or as part of the host chromosome. They also include cells which transiently express the inserted DNA or RNA for limited periods of time.

A "variant" of HRM refers to an amino acid sequence that is altered by one or more
30 amino acids. The variant may have "conservative" changes, wherein a substituted amino acid has similar structural or chemical properties, e.g., replacement of leucine with isoleucine. More rarely, a variant may have "nonconservative" changes, e.g., replacement of a glycine

with a tryptophan. Analogous minor variations may also include amino acid deletions or insertions, or both. Guidance in determining which amino acid residues may be substituted, inserted, or deleted without abolishing biological or immunological activity may be found using computer programs well known in the art, for example, LASERGENE software

5 (DNASTAR Inc., Madison WI)

THE INVENTION

The invention is based on the discovery of human regulatory molecules (HRM) and the polynucleotides encoding HRM and on the use of these compositions for the diagnosis, prevention, or treatment of diseases associated with cell proliferation. Table 1 shows the
10 protein and nucleotide sequence identification numbers, protein abbreviation, Incyte Clone number, cDNA library, NCBI homolog and NCBI sequence identifier for each of the human regulatory molecules disclosed herein.

5 HRM-1 (SEQ ID NO:1) was identified in Incyte Clone 9476 from the HMC1NOT01 cDNA library using a computer search for amino acid sequence alignments. A consensus
15 sequence, SEQ ID NO:11, was derived from the extended and overlapping nucleic acid sequences: Incyte Clones 9476 (HMC1NOT01), 010403 (THP1PLB01), 495099 (HNT2NOT01), 1670783 (BMARNOT03), 1997203 (BRSTTUT03), and 2190637 (THYRTUT03).

In one embodiment, the invention encompasses a polypeptide comprising the amino
20 acid sequence of SEQ ID NO:1. HRM-1 is 456 amino acids in length and has a potential N glycosylation site at N385; 14 potential phosphorylation sites at T9, T12, S58, T74, T163, T139, S175, T211, T239, T272, S331, T367, T402, and S443; and an ATP/GTP binding motif at G₇₀PPGTGKT77. HRM-1 has sequence homology with a *S. cerevisiae* protein (g755784) and is found in cDNA libraries which have dividing, cancerous or immortalized
25 cells and are associated with immune response.

8 HRM-2 (SEQ ID NO:2) was identified in Incyte Clone 77180 from the SYNORAB01 cDNA library using a computer search for amino acid sequence alignments. A consensus
sequence, SEQ ID NO:12, was derived from the extended and overlapping nucleic acid sequences: Incyte Clones 077180 (SYNORAB01), 604706 (BRSTTUT01), 977901
30 (BRSTNOT02), 1870373 (SKINBIT01), and 2169441 (ENDCNOT03).

In one embodiment, the invention encompasses a polypeptide comprising the amino acid sequence of SEQ ID NO:2. HRM-2 is 188 amino acids in length and has one potential

amidation site, Q170GKR; two potential N glycosylation sites at N60 and N68; and four potential phosphorylation sites at S70, T164, T166, and S183. HRM-2 has sequence homology with a *S. cerevisiae* protein (g5372) and is found in cDNA libraries which have dividing, cancerous or immortalized cells and are associated with immune response.

5 3 HRM-3 (SEQ ID NO:3) was identified in Incyte Clone 118160 from the MUSCNOT01 cDNA library using a computer search for amino acid sequence alignments. A consensus sequence, SEQ ID NO:13, was derived from the extended and overlapping nucleic acid sequences: Incyte Clones 118160 (MUSCNOT01), 323015 (EOSIHET02), and 1856519 (PROSNOT18).

10 In one embodiment, the invention encompasses a polypeptide comprising the amino acid sequence of SEQ ID NO:3. HRM-3 is 348 amino acids in length and has two potential N glycosylation sites at N150 and N317; 17 potential phosphorylation sites at T23, T45, S60, T126, S130, S140, S145, S151, S154, S158, S186, Y208, Y234, S217, T271, T303, and S327, and a transcription factor signature at C₃₁₀SKCKKKNCTYNQVQTRSA
15 DEPMTTFVLCNEC. HRM-3 has sequence homology with a *Mus musculus* protein (g220594) and is found in cDNA libraries which have secretory or immune associations.

21 HRM-4 (SEQ ID NO:4) was identified in Incyte Clone 933353 from the CERVNOT01 cDNA library using a computer search for amino acid sequence alignments. A consensus sequence, SEQ ID NO:14, was derived from the extended and overlapping nucleic
20 acid sequences: Incyte Clones 928904 (BRAINOT04), 933353 (CERVNOT01), and 2452674 (ENDANOT01).

 In one embodiment, the invention encompasses a polypeptide comprising the amino acid sequence of SEQ ID NO:4. HRM-4 is 201 amino acids in length and has one potential N glycosylation site at N82; five potential phosphorylation sites at T70, S83, S98, S154, and
25 T187; and one tyrosine phosphatase motif at V₁₃₀HCKAGRSRSATM. HRM-4 has sequence homology with a *C. elegans* protein (g1657672) and is found in cDNA libraries associated with immune response.

24 HRM-5 (SEQ ID NO:5) was identified in Incyte Clone 1404643 from the LATRTUT02 cDNA library using a computer search for amino acid sequence alignments. A
30 consensus sequence, SEQ ID NO:15, was derived from the extended and overlapping nucleic acid sequences: Incyte Clones 878243 (LUNGAST01), 1404643 (LATRTUT02), 1508343 (LUNGNOT14) and 2585156 (BRAITUT22).

In one embodiment, the invention encompasses a polypeptide comprising the amino acid sequence of SEQ ID NO:5. HRM-5 is 239 amino acids in length and has four potential phosphorylation sites at S5, S89, S133, and T211. HRM-5 has sequence homology with a C. elegans protein (g459002) and is found in cDNA libraries associated with cell proliferation,
5 cancer and immune response.

30 HRM-6 (SEQ ID NO:6) was identified in Incyte Clone 1600438 from the BLADNOT03 cDNA library using a computer search for amino acid sequence alignments. A consensus sequence, SEQ ID NO:16, was derived from the extended and overlapping nucleic acid sequences: Incyte Clones 835283 (PROSNOT07), 1600044 (BLADNOT03), 1600438
10 (BLADNOT03), and 1922072 (BRSTTUT01).

In one embodiment, the invention encompasses a polypeptide comprising the amino acid sequence of SEQ ID NO:6. HRM-6 is 419 amino acids in length and has one potential N glycosylation site at N161; twelve potential phosphorylation sites at T16, S57, T67, T83, S100, T107, S144, S206, T254, Y351, S412, and S414; a leucine zipper motif,
15 L₃₈NEAGDDLEAVAKFLDSTGSRL; and an ATP/GTP binding motif, A₃₈₅HVAKGKS. HRM-6 has sequence homology with human KIAA0005 (g286001) and is found in cDNA libraries associated with secretion and cancer.

33 HRM-7 (SEQ ID NO:7) was identified in Incyte Clone 1600518 from the BLADNOT03 cDNA library using a computer search for amino acid sequence alignments. A
20 consensus sequence, SEQ ID NO:17, was derived from the extended and overlapping nucleic acid sequences: Incyte Clones 389679 (THYMNOT02), 1600518 (BLADNOT03), 2055734 (BEPINOT01), 2102793 (BRAITUT02), and 2509270 (CONUTUT01).

In one embodiment, the invention encompasses a polypeptide comprising the amino acid sequence of SEQ ID NO:7. HRM-7 is 376 amino acids in length and has one potential N
25 glycosylation site at N161 and 14 potential phosphorylation sites at T30, S65, S75, S95, S106, T134, S159, S224, T228, T250, T292, S299, T303, and S323 and a glycosaminoglycan motif, S14GPG. HRM-7 has sequence homology with a C. elegans protein (g790405) and is found in cDNA libraries associated with immune response, secretion, and cancer.

45 HRM-8 (SEQ ID NO:8) was identified in Incyte Clone 2515476 from the LIVRTUT04 cDNA library using a computer search for amino acid sequence alignments. A
30 consensus sequence, SEQ ID NO:18, was derived from the extended and overlapping nucleic acid sequences: Incyte clones 18414 (HUVELPB01), 78341 (SYNORAB01), 143277

(TLYMNOR01), 181574 (PLACNOB01), 832996 (PROSTUT04), 962753 (BRSTTUT03), 1413604 (BRAINOT12), and 2515476 (LIVRTUT04).

In one embodiment, the invention encompasses a polypeptide comprising the amino acid sequence of SEQ ID NO:8. HRM-8 is 811 amino acids in length and has three potential
 5 amidation sites at G₁₁₃GRR, W₁₆₅GKR, and G₇₉₀GKK; four potential N glycosylation sites at N22, N56, N79, and N145; 24 potential phosphorylation sites at T11, S13, S30, S60, Y71, S81, S85, S86, S103, S254, S256, T377, S388, S425, S456, S487, T544, S552, S574, T659, S678, S702, S746, and S753; a potential glycosaminoglycan site, S₁₆₀GHG; and a potential zinc finger motif at C₂₄₀GHIFCWACI. HRM-8 has sequence homology with human
 10 KIAA0262 (g1665790) and is found in cDNA libraries involved in cell proliferation, secretion, cancer, and immune response.

46 HRM-9 (SEQ ID NO:9) was identified in Incyte Clone 2754573 from the THP1AZS08 cDNA library using a computer search for amino acid sequence alignments. A consensus sequence, SEQ ID NO:19, was derived from the extended and overlapping nucleic
 15 acid sequences: Incyte Clones 263630 (HNT2AGT01), 412307 (BRSTNOT01), 491644 (HNT2AGT01), 1253094 (LUNGFET03), 2270603 (PROSNON01), 2280508 (PROSNON01), 2375670 (ISLTNOT01), 2754573 (THP1AZS08), and 3151587 (ADRENON04).

In one embodiment, the invention encompasses a polypeptide comprising the amino
 20 acid sequence of SEQ ID NO:9. HRM-9 is 352 amino acids in length and has two potential N glycosylation sites at N141 and N294, and thirteen potential phosphorylation sites at S8, T67, T106, T110, T121, S122, S169, S206, T210, S215, S256, S260, and T296. HRM-9 has sequence homology with human RNA binding protein (g478990) and is found in cDNA libraries involved in cell proliferation, secretion, and immune response.

47 HRM-10 (SEQ ID NO:10) was identified in Incyte Clone 2926777 from the TLYMNOT04 cDNA library using a computer search for amino acid sequence alignments. A consensus sequence, SEQ ID NO:20, was derived from the extended and overlapping nucleic
 25 acid sequences: Incyte Clones 040208 (TBLYNOT01), 900242 (BRSTTUT03), 963500 (BRSTTUT03), 1996474 (BRSTTUT03), and 2926777 (TLYMNOT04).

30 In one embodiment, the invention encompasses a polypeptide comprising the amino acid sequence of SEQ ID NO:10. HRM-10 is 432 amino acids in length and has a potential N glycosylation site at N417 and 24 potential phosphorylation sites at T51, S73, T122, T133,

S177, S206, T226, T238, S293, S300, S304, S309, T325, S333, S339, S353, S360, Y361, S384, S390, T403, T412, T419, and S425. HRM-10 has sequence homology with a C. elegans protein (g687823) and is found in cDNA libraries involved in cell proliferation, secretion, cancer, and immune response.

5 The invention also encompasses HRM variants which retain the biological or functional activity of HRM. A preferred HRM variant is one having at least 80%, and more preferably 90%, amino acid sequence identity to the HRM amino acid sequence. A most preferred HRM variant is one having at least 95% amino acid sequence identity to an HRM disclosed herein (SEQ ID NOs:1-10).

10 The invention also encompasses polynucleotides which encode HRM. Accordingly, any nucleic acid sequence which encodes the amino acid sequence of HRM can be used to produce recombinant molecules which express HRM. In a particular embodiment, the invention encompasses a polynucleotide consisting of a nucleic acid sequence selected from the group consisting of SEQ ID NOs:11-20.

15 It will be appreciated by those skilled in the art that as a result of the degeneracy of the genetic code, a multitude of nucleotide sequences encoding HRM, some bearing minimal homology to the nucleotide sequences of any known and naturally occurring gene, may be produced. Thus, the invention contemplates each and every possible variation of nucleotide sequence that could be made by selecting combinations based on possible codon choices.
20 These combinations are made in accordance with the standard triplet genetic code as applied to the nucleotide sequence of naturally occurring HRM, and all such variations are to be considered as being specifically disclosed.

 Although nucleotide sequences which encode HRM and its variants are preferably capable of hybridizing to the nucleotide sequence of the naturally occurring HRM under
25 appropriately selected conditions of stringency, it may be advantageous to produce nucleotide sequences encoding HRM or its derivatives possessing a substantially different codon usage. Codons may be selected to increase the rate at which expression of the peptide occurs in a particular prokaryotic or eukaryotic host in accordance with the frequency with which particular codons are utilized by the host. Other reasons for substantially altering the
30 nucleotide sequence encoding HRM and its derivatives without altering the encoded amino acid sequences include the production of RNA transcripts having more desirable properties, such as a greater half-life, than transcripts produced from the naturally occurring sequence.

The invention also encompasses production of DNA sequences, or fragments thereof, which encode HRM and its derivatives, entirely by synthetic chemistry. After production, the synthetic sequence may be inserted into any of the many available expression vectors and cell systems using reagents that are well known in the art. Moreover, synthetic chemistry may be used to introduce mutations into a sequence encoding HRM or any fragment thereof.

Also encompassed by the invention are polynucleotide sequences that are capable of hybridizing to the claimed nucleotide sequences, and in particular, those shown in SEQ ID NOs:11-20, under various conditions of stringency as taught in Wahl, G.M. and S.L. Berger (1987; Methods Enzymol. 152:399-407) and Kimmel, A.R. (1987; Methods Enzymol. 152:507-511).

Methods for DNA sequencing which are well known and generally available in the art and may be used to practice any of the embodiments of the invention. The methods may employ such enzymes as the Klenow fragment of DNA polymerase I, SEQUENASE (US Biochemical Corp, Cleveland OH), TAQ polymerase (Perkin Elmer), thermostable T7 polymerase (Amersham Pharmacia, Piscataway NJ), or combinations of polymerases and proofreading exonucleases such as those found in the ELONGASE amplification system marketed by Life Technologies (Gaithersburg MD). Preferably, the process is automated with machines such as the Hamilton Micro Lab 2200 (Hamilton, Reno NV), Peltier Thermal Cycler (PTC200; MJ Research, Watertown MA) and the ABI Catalyst and 373 and 377 DNA Sequencers (Perkin Elmer).

The nucleic acid sequences encoding HRM may be extended utilizing a partial nucleotide sequence and employing various methods known in the art to detect upstream sequences such as promoters and regulatory elements. For example, one method which may be employed, "restriction-site" PCR, uses universal primers to retrieve unknown sequence adjacent to a known locus (Sarkar, G. (1993) PCR Methods Applic. 2:318-322). In particular, genomic DNA is first amplified in the presence of primer to a linker sequence and a primer specific to the known region. The amplified sequences are then subjected to a second round of PCR with the same linker primer and another specific primer internal to the first one. Products of each round of PCR are transcribed with an appropriate RNA polymerase and sequenced using reverse transcriptase.

Inverse PCR may also be used to amplify or extend sequences using divergent primers based on a known region (Triglia, T. et al. (1988) Nucleic Acids Res. 16:8186). The primers

may be designed using commercially available software such as OLIGO 4.06 Primer Analysis software (National Biosciences Inc., Plymouth MN), or another appropriate program, to be 22-30 nucleotides in length, to have a GC content of 50% or more, and to anneal to the target sequence at temperatures about 68°-72° C. The method uses several restriction enzymes to
5 generate a suitable fragment in the known region of a gene. The fragment is then circularized by intramolecular ligation and used as a PCR template.

Another method which may be used is capture PCR which involves PCR amplification of DNA fragments adjacent to a known sequence in human and yeast artificial chromosome DNA (Lagerstrom, M. et al. (1991) PCR Methods Applic. 1:111-119). In this
10 method, multiple restriction enzyme digestions and ligations may also be used to place an engineered double-stranded sequence into an unknown fragment of the DNA molecule before performing PCR.

Another method which may be used to retrieve unknown sequences is that of Parker, J.D. et al. (1991: Nucleic Acids Res. 19:3055-3060). Additionally, one may use PCR, nested
15 primers, and PROMOTERFINDER libraries to walk genomic DNA (Clontech, Palo Alto CA). This process avoids the need to screen libraries and is useful in finding intron/exon junctions. When screening for full-length cDNAs, it is preferable to use libraries that have been size-selected to include larger cDNAs. Also, random-primed libraries are preferable, in that they will contain more sequences which contain the 5' regions of genes. Use of a
20 randomly primed library may be especially preferable for situations in which an oligo d(T) library does not yield a full-length cDNA. Genomic libraries may be useful for extension of sequence into 5' non-transcribed regulatory regions.

Capillary electrophoresis systems which are commercially available may be used to analyze the size or confirm the nucleotide sequence of sequencing or PCR products. In
25 particular, capillary sequencing may employ flowable polymers for electrophoretic separation, four different fluorescent dyes (one for each nucleotide) which are laser activated, and detection of the emitted wavelengths by a charge coupled device camera. Output/light intensity may be converted to electrical signal using appropriate software (e.g., GENOTYPER and SEQUENCE NAVIGATOR, Perkin Elmer) and the entire process from loading of
30 samples to computer analysis and electronic data display may be computer controlled. Capillary electrophoresis is especially preferable for the sequencing of small pieces of DNA which might be present in limited amounts in a particular sample.

In another embodiment of the invention, polynucleotide sequences or fragments thereof which encode HRM may be used in recombinant DNA molecules to direct expression of HRM, fragments or functional equivalents thereof, in appropriate host cells. Due to the inherent degeneracy of the genetic code, other DNA sequences which encode substantially the same or a functionally equivalent amino acid sequence may be produced, and these sequences
5 may be used to clone and express HRM.

As will be understood by those of skill in the art, it may be advantageous to produce HRM-encoding nucleotide sequences possessing non-naturally occurring codons. For example, codons preferred by a particular prokaryotic or eukaryotic host can be selected to
10 increase the rate of protein expression or to produce an RNA transcript having desirable properties, such as a half-life which is longer than that of a transcript generated from the naturally occurring sequence.

The nucleotide sequences of the present invention can be engineered using methods generally known in the art in order to alter HRM encoding sequences for a variety of reasons, including but not limited to, alterations which modify the cloning, processing, and/or
15 expression of the gene product. DNA shuffling by random fragmentation and PCR reassembly of gene fragments and synthetic oligonucleotides may be used to engineer the nucleotide sequences. For example, site-directed mutagenesis may be used to insert new restriction sites, alter glycosylation patterns, change codon preference, produce splice
20 variants, introduce mutations, and so forth.

In another embodiment of the invention, natural, modified, or recombinant nucleic acid sequences encoding HRM may be ligated to a heterologous sequence to encode a fusion protein. For example, to screen peptide libraries for inhibitors of HRM activity, it may be useful to encode a chimeric HRM protein that can be recognized by a commercially available
25 antibody. A fusion protein may also be engineered to contain a cleavage site located between the HRM encoding sequence and the heterologous protein sequence, so that HRM may be cleaved and purified away from the heterologous moiety.

In another embodiment, sequences encoding HRM may be synthesized, in whole or in part, using chemical methods well known in the art. (See Caruthers, M.H. et al. (1980) Nucl. Acids Res. Symp. Ser. 215-223; Horn, T. et al. (1980) Nucl. Acids Res. Symp. Ser. 225-232.)
30 Alternatively, the protein itself may be produced using chemical methods to synthesize the amino acid sequence of HRM, or a fragment thereof. For example, peptide synthesis can be

performed using various solid-phase techniques (Roberge, J.Y. et al. (1995) Science 269:202-204) and automated synthesis may be achieved, for example, using the ABI 431A Peptide synthesizer (Perkin Elmer).

5 The newly synthesized peptide may be substantially purified by preparative high performance liquid chromatography (e.g., Creighton, T. (1983) Proteins, Structures and Molecular Principles, WH Freeman and Co., New York NY). The composition of the synthetic peptides may be confirmed by amino acid analysis or sequencing (e.g., the Edman degradation procedure; Creighton, supra). Additionally, the amino acid sequence of HRM, or any part thereof, may be altered during direct synthesis and/or combined using chemical
10 methods with sequences from other proteins, or any part thereof, to produce a variant polypeptide.

In order to express a biologically active HRM, the nucleotide sequences encoding HRM or functional equivalents, may be inserted into appropriate expression vector, i.e., a vector which contains the necessary elements for the transcription and translation of the
15 inserted coding sequence.

Methods which are well known to those skilled in the art may be used to construct expression vectors containing sequences encoding HRM and appropriate transcriptional and translational control elements. These methods include in vitro recombinant DNA techniques, synthetic techniques, and in vivo genetic recombination. Such techniques are described in
20 Sambrook, J. et al. (1989; Molecular Cloning, A Laboratory Manual, Cold Spring Harbor Press, Plainview NY) and Ausubel, F.M. et al. (1989; Current Protocols in Molecular Biology, John Wiley & Sons, New York NY).

A variety of expression vector/host systems may be utilized to contain and express sequences encoding HRM. These include, but are not limited to, microorganisms such as
25 bacteria transformed with recombinant bacteriophage, plasmid, or cosmid DNA expression vectors; yeast transformed with yeast expression vectors; insect cell systems infected with virus expression vectors (e.g., baculovirus); plant cell systems transformed with virus expression vectors (e.g., cauliflower mosaic virus, CaMV; tobacco mosaic virus, TMV) or with bacterial expression vectors (e.g., Ti or pBR322 plasmids); or animal cell systems. The
30 invention is not limited by the host cell employed.

The "control elements" or "regulatory sequences" are those non-translated regions of the vector--enhancers, promoters, 5' and 3' untranslated regions--which interact with host

cellular proteins to carry out transcription and translation. Such elements may vary in their strength and specificity. Depending on the vector system and host utilized, any number of suitable transcription and translation elements, including constitutive and inducible promoters, may be used. For example, when cloning in bacterial systems, inducible
5 promoters such as the hybrid lacZ promoter of the BLUESCRIPT phagemid (Stratagene, LaJolla CA) or pSPORT1 plasmid (Life Technologies) and the like may be used. The baculovirus polyhedrin promoter may be used in insect cells. Promoters or enhancers derived from the genomes of plant cells (e.g., heat shock, RUBISCO; and storage protein genes) or from plant viruses (e.g., viral promoters or leader sequences) may be cloned into the vector.
10 In mammalian cell systems, promoters from mammalian genes or from mammalian viruses are preferable. If it is necessary to generate a cell line that contains multiple copies of the sequence encoding HRM, vectors based on SV40 or EBV may be used with an appropriate selectable marker.

In bacterial systems, a number of expression vectors may be selected depending upon
15 the use intended for HRM. For example, when large quantities of HRM are needed for the induction of antibodies, vectors which direct high level expression of fusion proteins that are readily purified may be used. Such vectors include, but are not limited to, the multifunctional *E. coli* cloning and expression vectors such as BLUESCRIPT (Stratagene), in which the sequence encoding HRM may be ligated into the vector in frame with sequences for the
20 amino-terminal Met and the subsequent 7 residues of β -galactosidase so that a hybrid protein is produced; pIN vectors (Van Heeke, G. and S.M. Schuster (1989) J. Biol. Chem. 264:5503-5509); and the like. pGEX vectors (Promega, Madison WI) may also be used to express foreign polypeptides as fusion proteins with glutathione S-transferase (GST). In general, such fusion proteins are soluble and can easily be purified from lysed cells by
25 adsorption to glutathione-agarose beads followed by elution in the presence of free glutathione. Proteins made in such systems may be designed to include heparin, thrombin, or factor XA protease cleavage sites so that the cloned polypeptide of interest can be released from the GST moiety at will.

In the yeast, *Saccharomyces cerevisiae*, a number of vectors containing constitutive or
30 inducible promoters such as alpha factor, alcohol oxidase, and PGH may be used. For reviews, see Ausubel et al. (supra) and Grant et al. (1987: Methods Enzymol. 153:516-544).

In cases where plant expression vectors are used, the expression of sequences

encoding HRM may be driven by any of a number of promoters. For example, viral promoters such as the 35S and 19S promoters of CaMV may be used alone or in combination with the omega leader sequence from TMV (Takamatsu, N. (1987) EMBO J. 6:307-311). Alternatively, plant promoters such as the small subunit of RUBISCO or heat shock promoters may be used (Coruzzi, G. et al. (1984) EMBO J. 3:1671-1680; Broglie, R. et al. (1984) Science 224:838-843; and Winter, J. et al. (1991) Results Probl. Cell Differ. 17:85-105). These constructs can be introduced into plant cells by direct DNA transformation or pathogen-mediated transfection. Such techniques are described in a number of generally available reviews (McGraw Hill Yearbook of Science and Technology (1992) McGraw Hill, New York NY; pp. 191-196).

An insect system may also be used to express HRM. For example, in one such system, Autographa californica nuclear polyhedrosis virus (AcNPV) is used as a vector to express foreign genes in Spodoptera frugiperda cells or in Trichoplusia larvae. The sequences encoding HRM may be cloned into a non-essential region of the virus, such as the polyhedrin gene, and placed under control of the polyhedrin promoter. Successful insertion of HRM will render the polyhedrin gene inactive and produce recombinant virus lacking coat protein. The recombinant viruses may then be used to infect, for example, S. frugiperda cells or Trichoplusia larvae in which HRM may be expressed (Engelhard, E.K. et al. (1994) Proc. Nat. Acad. Sci. 91:3224-3227).

In mammalian host cells, a number of viral-based expression systems may be utilized. In cases where an adenovirus is used as an expression vector, sequences encoding HRM may be ligated into an adenovirus transcription/translation complex consisting of the late promoter and tripartite leader sequence. Insertion in a non-essential E1 or E3 region of the viral genome may be used to obtain a viable virus which is capable of expressing HRM in infected host cells (Logan, J. and Shenk, T. (1984) Proc. Natl. Acad. Sci. 81:3655-3659). In addition, transcription enhancers, such as the Rous sarcoma virus (RSV) enhancer, may be used to increase expression in mammalian host cells.

Human artificial chromosomes (HACs) may also be employed to deliver larger fragments of DNA than can be contained and expressed in a plasmid. HACs of 6 to 10M are constructed and delivered via conventional delivery methods (liposomes, polycationic amino polymers, or vesicles) for therapeutic purposes (Harrington, J.J. et al. (1997) Nat Genet. 15:345-355).

Specific initiation signals may also be used to achieve more efficient translation of sequences encoding HRM. Such signals include the ATG initiation codon and adjacent sequences. In cases where sequences encoding HRM, its initiation codon, and upstream sequences are inserted into the appropriate expression vector, no additional transcriptional or translational control signals may be needed. However, in cases where only coding sequence, or a fragment thereof, is inserted, exogenous translational control signals including the ATG initiation codon should be provided. Furthermore, the initiation codon should be in the correct reading frame to ensure translation of the entire insert. Exogenous translational elements and initiation codons may be of various origins, both natural and synthetic. The efficiency of expression may be enhanced by the inclusion of enhancers which are appropriate for the particular cell system which is used, such as those described in the literature (Scharf, D. et al. (1994) *Results Probl. Cell Differ.* 20:125-162).

In addition, a host cell strain may be chosen for its ability to modulate the expression of the inserted sequences or to process the expressed protein in the desired fashion. Such modifications of the polypeptide include, but are not limited to, acetylation, carboxylation, glycosylation, phosphorylation, lipidation, and acylation. Post-translational processing which cleaves a "prepro" form of the protein may also be used to facilitate correct insertion, folding and/or function. Different host cells which have specific cellular machinery and characteristic mechanisms for post-translational activities (e.g., CHO, HeLa, MDCK, HEK293, and WI38), are available from the American Type Culture Collection (ATCC; Bethesda MD) and may be chosen to ensure the correct modification and processing of the foreign protein.

For long-term, high-yield production of recombinant proteins, stable expression is preferred. For example, cell lines which stably express HRM may be transformed using expression vectors which may contain viral origins of replication and/or endogenous expression elements and a selectable marker gene on the same or on a separate vector. Following the introduction of the vector, cells may be allowed to grow for 1-2 days in an enriched media before they are switched to selective media. The purpose of the selectable marker is to confer resistance to selection, and its presence allows growth and recovery of cells which successfully express the introduced sequences. Resistant clones of stably transformed cells may be proliferated using tissue culture techniques appropriate to the cell type.

Any number of selection systems may be used to recover transformed cell lines.

These include, but are not limited to, the herpes simplex virus thymidine kinase (Wigler, M. et al. (1977) *Cell* 11:223-32) and adenine phosphoribosyltransferase (Lowy, I. et al. (1980) *Cell* 22:817-23) genes which can be employed in tk⁻ or apt⁻ cells, respectively. Also, antimetabolite, antibiotic or herbicide resistance can be used as the basis for selection; for example, dhfr which confers resistance to methotrexate (Wigler, M. et al. (1980) *Proc. Natl. Acad. Sci.* 77:3567-70); npt, which confers resistance to the aminoglycosides neomycin and G-418 (Colbere-Garapin, F. et al (1981) *J. Mol. Biol.* 150:1-14) and als or pat, which confer resistance to chlorsulfuron and phosphinotricin acetyltransferase, respectively (McGraw Hill Yearbook of Science and Technology, supra). Additional selectable genes have been described, for example, trpB, which allows cells to utilize indole in place of tryptophan, or hisD, which allows cells to utilize histinol in place of histidine (Hartman, S.C. and R.C. Mulligan (1988) *Proc. Natl. Acad. Sci.* 85:8047-51). Recently, the use of visible markers has gained popularity with such markers as anthocyanins, β glucuronidase and its substrate GUS, and luciferase and its substrate luciferin, being widely used not only to identify transformants, but also to quantify the amount of transient or stable protein expression attributable to a specific vector system (Rhodes, C.A. et al. (1995) *Methods Mol. Biol.* 55:121-131).

Although the presence/absence of marker gene expression suggests that the gene of interest is also present, its presence and expression may need to be confirmed. For example, if the sequence encoding HRM is inserted within a marker gene sequence, transformed cells containing sequences encoding HRM can be identified by the absence of marker gene function. Alternatively, a marker gene can be placed in tandem with a sequence encoding HRM under the control of a single promoter. Expression of the marker gene in response to induction or selection usually indicates expression of the tandem gene as well.

Alternatively, host cells which contain the nucleic acid sequence encoding HRM and express HRM may be identified by a variety of procedures known to those of skill in the art. These procedures include, but are not limited to, DNA-DNA or DNA-RNA hybridizations and protein bioassay or immunoassay techniques which include membrane, solution, or chip based technologies for the detection and/or quantification of nucleic acid or protein.

The presence of polynucleotide sequences encoding HRM can be detected by DNA-DNA or DNA-RNA hybridization or amplification using probes or fragments or fragments of polynucleotides encoding HRM. Nucleic acid amplification based assays involve the use of oligonucleotides or oligomers based on the sequences encoding HRM to

detect transformants containing DNA or RNA encoding HRM.

A variety of protocols for detecting and measuring the expression of HRM, using either polyclonal or monoclonal antibodies specific for the protein are known in the art. Examples include enzyme-linked immunosorbent assay (ELISA), radioimmunoassay (RIA),
5 and fluorescence activated cell sorting (FACS). A two-site, monoclonal-based immunoassay utilizing monoclonal antibodies reactive to two non-interfering epitopes on HRM is preferred, but a competitive binding assay may be employed. These and other assays are described, among other places, in Hampton, R. et al. (1990; Serological Methods, a Laboratory Manual, APS Press, St Paul MN) and Maddox, D.E. et al. (1983; J. Exp. Med. 158:1211-1216).

10 A wide variety of labels and conjugation techniques are known by those skilled in the art and may be used in various nucleic acid and amino acid assays. Means for producing labeled hybridization or PCR probes for detecting sequences related to polynucleotides encoding HRM include oligolabeling, nick translation, end-labeling or PCR amplification using a labeled nucleotide. Alternatively, the sequences encoding HRM, or any fragments
15 thereof may be cloned into a vector for the production of an mRNA probe. Such vectors are known in the art, are commercially available, and may be used to synthesize RNA probes in vitro by addition of an appropriate RNA polymerase such as T7, T3, or SP6 and labeled nucleotides. These procedures may be conducted using a variety of commercially available kits (Amersham Pharmacia, Promega; and US Biochemical Corp.). Suitable reporter
20 molecules or labels, which may be used for ease of detection, include radionuclides, enzymes, fluorescent, chemiluminescent, or chromogenic agents as well as substrates, cofactors, inhibitors, magnetic particles, and the like.

Host cells transformed with nucleotide sequences encoding HRM may be cultured under conditions suitable for the expression and recovery of the protein from cell culture.
25 The protein produced by a transformed cell may be secreted or contained intracellularly depending on the sequence and/or the vector used. As will be understood by those of skill in the art, expression vectors containing polynucleotides which encode HRM may be designed to contain signal sequences which direct secretion of HRM through a prokaryotic or eukaryotic cell membrane. Other constructions may be used to join sequences encoding
30 HRM to nucleotide sequence encoding a polypeptide domain which will facilitate purification of soluble proteins. Such purification facilitating domains include, but are not limited to, metal chelating peptides such as histidine-tryptophan modules that allow purification on

immobilized metals, protein A domains that allow purification on immobilized immunoglobulin, and the domain utilized in the FLAGS extension/affinity purification system (Immunex Corp., Seattle WA). The inclusion of cleavable linker sequences such as those specific for Factor XA or enterokinase (Invitrogen, San Diego CA) between the purification domain and HRM may be used to facilitate purification. One such expression vector provides for expression of a fusion protein containing HRM and a nucleic acid encoding 6 histidine residues preceding a thioredoxin or an enterokinase cleavage site. The histidine residues facilitate purification on IMAC (immobilized metal ion affinity chromatography as described in Porath, J. et al. (1992, Prot. Exp. Purif. 3: 263-281) while the enterokinase cleavage site provides a means for purifying HRM from the fusion protein. A discussion of vectors which contain fusion proteins is provided in Kroll, D.J. et al. (1993; DNA Cell Biol. 12:441-453).

In addition to recombinant production, fragments of HRM may be produced by direct peptide synthesis using solid-phase techniques (Merrifield J. (1963) J. Am. Chem. Soc. 85:2149-2154). Protein synthesis may be performed using manual techniques or by automation. Automated synthesis may be achieved, for example, using Applied Biosystems 431A Peptide synthesizer (Perkin Elmer). Various fragments of HRM may be chemically synthesized separately and combined using chemical methods to produce the full length molecule.

20 THERAPEUTICS

Chemical and structural homology exists among the human regulatory proteins of the invention. The expression of HRM is closely associated with cell proliferation. Therefore, in cancers or immune disorders where HRM is an activator, transcription factor, or enhancer, and is promoting cell proliferation; it is desirable to decrease the expression of HRM. In cancers where HRM is an inhibitor or suppressor and is controlling or decreasing cell proliferation, it is desirable to provide the protein or to increase the expression of HRM.

In one embodiment, where HRM is an inhibitor, HRM or a fragment or derivative thereof may be administered to a subject to prevent or treat a cancer such as adenocarcinoma, leukemia, lymphoma, melanoma, myeloma, sarcoma, and teratocarcinoma. Such cancers include, but are not limited to, cancers of the adrenal gland, bladder, bone, bone marrow, brain, breast, cervix, gall bladder, ganglia, gastrointestinal tract, heart, kidney, liver, lung, muscle, ovary, pancreas, parathyroid, penis, prostate, salivary glands, skin, spleen, testis,

thymus, thyroid, and uterus.

In another embodiment, an agonist which is specific for HRM may be administered to a subject to prevent or treat a cancer including, but not limited to, those cancers listed above.

In another further embodiment, a vector capable of expressing HRM, or a fragment or
5 a derivative thereof, may be administered to a subject to prevent or treat a cancer including, but not limited to, those cancers listed above.

In a further embodiment where HRM is promoting cell proliferation, antagonists which decrease the expression or activity of HRM may be administered to a subject to prevent or treat a cancer such as adenocarcinoma, leukemia, lymphoma, melanoma, myeloma,
10 sarcoma, and teratocarcinoma. Such cancers include, but are not limited to, cancers of the adrenal gland, bladder, bone, bone marrow, brain, breast, cervix, gall bladder, ganglia, gastrointestinal tract, heart, kidney, liver, lung, muscle, ovary, pancreas, parathyroid, penis, prostate, salivary glands, skin, spleen, testis, thymus, thyroid, and uterus. In one aspect, antibodies which specifically bind HRM may be used directly as an antagonist or indirectly as
15 a targeting or delivery mechanism for bringing a pharmaceutical agent to cells or tissue which express HRM.

In another embodiment, a vector expressing the complement of the polynucleotide encoding HRM may be administered to a subject to treat or prevent a cancer including, but not limited to, those cancers listed above.

In yet another embodiment where HRM is promoting leukocyte activity or proliferation, antagonists which decrease the activity of HRM may be administered to a subject to prevent or treat an immune response. Such responses may be associated with AIDS, Addison's disease, adult respiratory distress syndrome, allergies, anemia, asthma, atherosclerosis, bronchitis, cholecystitis, Crohn's disease, ulcerative colitis, atopic dermatitis,
25 dermatomyositis, diabetes mellitus, emphysema, atrophic gastritis, glomerulonephritis, gout, Graves' disease, hypereosinophilia, irritable bowel syndrome, lupus erythematosus, multiple sclerosis, myasthenia gravis, myocardial or pericardial inflammation, osteoarthritis, osteoporosis, pancreatitis, polymyositis, rheumatoid arthritis, scleroderma, Sjögren's syndrome, and autoimmune thyroiditis; complications of cancer, hemodialysis, extracorporeal
30 circulation; viral, bacterial, fungal, parasitic, protozoal, and helminthic infections; and trauma. In one aspect, antibodies which specifically bind HRM may be used directly as an antagonist or indirectly as a targeting or delivery mechanism for bringing a pharmaceutical

agent to cells or tissue which express HRM.

In another embodiment, a vector expressing the complement of the polynucleotide encoding HRM may be administered to a subject to treat or prevent an immune response including, but not limited to, those listed above

5 In one further embodiment, HRM or a fragment or derivative thereof may be added to cells to stimulate cell proliferation. In particular, HRM may be added to a cell in culture or cells in vivo using delivery mechanisms such as liposomes, viral based vectors, or electroinjection for the purpose of promoting cell proliferation and tissue or organ regeneration. Specifically, HRM may be added to a cell, cell line, tissue or organ culture in
10 vitro or ex vivo to stimulate cell proliferation for use in heterologous or autologous transplantation. In some cases, the cell will have been preselected for its ability to fight an infection or a cancer or to correct a genetic defect in a disease such as sickle cell anemia, β thalassemia, cystic fibrosis, or Huntington's chorea.

In another embodiment, an agonist which is specific for HRM may be administered to
15 a cell to stimulate cell proliferation, as described above.

In another embodiment, a vector capable of expressing HRM, or a fragment or a derivative thereof, may be administered to a cell to stimulate cell proliferation, as described above.

In other embodiments, any of the therapeutic proteins, antagonists, antibodies,
20 agonists, complementary sequences or vectors of the invention may be administered in combination with other appropriate therapeutic agents. Selection of the appropriate agents for use in combination therapy may be made by one of ordinary skill in the art, according to conventional pharmaceutical principles. The combination of therapeutic agents may act synergistically to effect the treatment or prevention of the various disorders described above.
25 Using this approach, one may be able to achieve therapeutic efficacy with lower dosages of each agent, thus reducing the potential for adverse side effects.

Antagonists or inhibitors of HRM may be produced using methods which are generally known in the art. In particular, purified HRM may be used to produce antibodies or to screen libraries of pharmaceutical agents to identify those which specifically bind HRM.

30 Antibodies to HRM may be generated using methods that are well known in the art. Such antibodies may include, but are not limited to, polyclonal, monoclonal, chimeric, single chain, Fab fragments, and fragments produced by a Fab expression library. Neutralizing

antibodies, (i.e., those which inhibit dimer formation) are especially preferred for therapeutic use.

For the production of antibodies, various hosts including goats, rabbits, rats, mice, humans, and others, may be immunized by injection with HRM or any fragment or
5 oligopeptide thereof which has immunogenic properties. Depending on the host species, various adjuvants may be used to increase immunological response. Such adjuvants include, but are not limited to, Freund's, mineral gels such as aluminum hydroxide, and surface active substances such as lysolecithin, pluronic polyols, polyanions, peptides, oil emulsions, keyhole limpet hemocyanin, and dinitrophenol. Among adjuvants used in humans, BCG (bacilli
10 Calmette-Guerin) and Corynebacterium parvum are especially preferable.

It is preferred that the oligopeptides, peptides, or fragments used to induce antibodies to HRM have an amino acid sequence consisting of at least five amino acids and more preferably at least 10 amino acids. It is also preferable that they are identical to a portion of the amino acid sequence of the natural protein, and they may contain the entire amino acid
15 sequence of a small, naturally occurring molecule. Short stretches of HRM amino acids may be fused with those of another protein such as keyhole limpet hemocyanin and antibody produced against the chimeric molecule.

Monoclonal antibodies to HRM may be prepared using any technique which provides for the production of antibody molecules by continuous cell lines in culture. These include,
20 but are not limited to, the hybridoma technique, the human B-cell hybridoma technique, and the EBV-hybridoma technique (Kohler, G. et al. (1975) Nature 256:495-497; Kozbor, D. et al. (1985) J. Immunol. Methods 81:31-42; Cote, R.J. et al. (1983) Proc. Natl. Acad. Sci. 80:2026-2030; and Cole, S.P. et al. (1984) Mol. Cell Biol. 62:109-120).

In addition, techniques developed for the production of "chimeric antibodies", the
25 splicing of mouse antibody genes to human antibody genes to obtain a molecule with appropriate antigen specificity and biological activity can be used (Morrison, S.L. et al. (1984) Proc. Natl. Acad. Sci. 81:6851-6855; Neuberger, M.S. et al. (1984) Nature 312:604-608; and Takeda, S. et al. (1985) Nature 314:452-454). Alternatively, techniques described for the production of single chain antibodies may be adapted, using methods known
30 in the art, to produce HRM-specific single chain antibodies. Antibodies with related specificity, but of distinct idiotypic composition, may be generated by chain shuffling from random combinatorial immunoglobulin libraries (Burton D.R. (1991) Proc. Natl. Acad. Sci.

88:11120-3).

Antibodies may also be produced by inducing *in vivo* production in the lymphocyte population or by screening immunoglobulin libraries or panels of highly specific binding reagents as disclosed in the literature (Orlandi, R. et al. (1989) Proc. Natl. Acad. Sci. 86: 3833-3837; Winter, G. et al. (1991) Nature 349:293-299).

Antibody fragments which contain specific binding sites for HRM may also be generated. For example, such fragments include, but are not limited to, the F(ab')₂ fragments which can be produced by pepsin digestion of the antibody molecule and the Fab fragments which can be generated by reducing the disulfide bridges of the F(ab')₂ fragments.

Alternatively, Fab expression libraries may be constructed to allow rapid and easy identification of monoclonal Fab fragments with the desired specificity (Huse, W.D. et al. (1989) Science 254:1275-1281).

Various immunoassays may be used for screening to identify antibodies having the desired specificity. Numerous protocols for competitive binding or immunoradiometric assays using either polyclonal or monoclonal antibodies with established specificities are well known in the art. Such immunoassays typically involve the measurement of complex formation between HRM and its specific antibody. A two-site, monoclonal-based immunoassay utilizing monoclonal antibodies reactive to two non-interfering HRM epitopes is preferred, but a competitive binding assay may also be employed (Maddox, supra).

In another embodiment of the invention, the polynucleotides encoding HRM, or any fragment or complement thereof, may be used for therapeutic purposes. In one aspect, the complement of the polynucleotide encoding HRM may be used in situations in which it would be desirable to block the transcription of the mRNA. In particular, cells may be transformed with sequences complementary to polynucleotides encoding HRM. Thus, complementary molecules or fragments may be used to modulate HRM activity, or to achieve regulation of gene function. Such technology is now well known in the art, and sense or antisense oligonucleotides or larger fragments, can be designed from various locations along the coding or control regions of sequences encoding HRM.

Expression vectors derived from retroviruses, adenovirus, herpes or vaccinia viruses, or from various bacterial plasmids may be used for delivery of nucleotide sequences to the targeted organ, tissue or cell population. Methods which are well known to those skilled in the art can be used to construct vectors which will express nucleic acid sequence which is

complementary to the polynucleotides of the gene encoding HRM. These techniques are described both in Sambrook et al. (supra) and in Ausubel et al. (supra).

Genes encoding HRM can be turned off by transforming a cell or tissue with expression vectors which express high levels of a polynucleotide or fragment thereof which encodes HRM. Such constructs may be used to introduce untranslatable sense or antisense sequences into a cell. Even in the absence of integration into the DNA, such vectors may continue to transcribe RNA molecules until they are disabled by endogenous nucleases. Transient expression may last for a month or more with a non-replicating vector and even longer if appropriate replication elements are part of the vector system.

As mentioned above, modifications of gene expression can be obtained by designing complementary sequences or antisense molecules (DNA, RNA, or PNA) to the control, 5' or regulatory regions of the gene encoding HRM (signal sequence, promoters, enhancers, and introns). Oligonucleotides derived from the transcription initiation site, e.g., between positions -10 and +10 from the start site, are preferred. Similarly, inhibition can be achieved using "triple helix" base-pairing methodology. Triple helix pairing is useful because it causes inhibition of the ability of the double helix to open sufficiently for the binding of polymerases, transcription factors, or regulatory molecules. Recent therapeutic advances using triplex DNA have been described in the literature (Gee, J.E. et al. (1994) In: Huber, B.E. and B.I. Carr, Molecular and Immunologic Approaches, Futura Publishing Co., Mt. Kisco, NY). The complementary sequence or antisense molecule may also be designed to block translation of mRNA by preventing the transcript from binding to ribosomes.

Ribozymes, enzymatic RNA molecules, may also be used to catalyze the specific cleavage of RNA. The mechanism of ribozyme action involves sequence-specific hybridization of the ribozyme molecule to complementary target RNA, followed by endonucleolytic cleavage. Examples which may be used include engineered hammerhead motif ribozyme molecules that can specifically and efficiently catalyze endonucleolytic cleavage of sequences encoding HRM.

Specific ribozyme cleavage sites within any potential RNA target are initially identified by scanning the target molecule for ribozyme cleavage sites which include the following sequences: GUA, GUU, and GUC. Once identified, short RNA sequences of between 15 and 20 ribonucleotides corresponding to the region of the target gene containing the cleavage site may be evaluated for secondary structural features which may render the

oligonucleotide inoperable. The suitability of candidate targets may also be evaluated by testing accessibility to hybridization with complementary oligonucleotides using ribonuclease protection assays.

Complementary ribonucleic acid molecules and ribozymes of the invention may be prepared by any method known in the art for the synthesis of nucleic acid molecules. These include techniques for chemically synthesizing oligonucleotides such as solid phase phosphoramidite chemical synthesis. Alternatively, RNA molecules may be generated by *in vitro* and *in vivo* transcription of DNA sequences encoding HRM. Such DNA sequences may be incorporated into a wide variety of vectors with suitable RNA polymerase promoters such as T7 or SP6. Alternatively, these cDNA constructs that synthesize complementary RNA constitutively or inducibly can be introduced into cell lines, cells, or tissues.

RNA molecules may be modified to increase intracellular stability and half-life. Possible modifications include, but are not limited to, the addition of flanking sequences at the 5' and/or 3' ends of the molecule or the use of phosphorothioate or 2' O-methyl rather than phosphodiesterase linkages within the backbone of the molecule. This concept is inherent in the production of PNAs and can be extended in all of these molecules by the inclusion of nontraditional bases such as inosine, queosine, and wybutosine, as well as acetyl-, methyl-, thio-, and similarly modified forms of adenine, cytidine, guanine, thymine, and uridine which are not as easily recognized by endogenous endonucleases.

Many methods for introducing vectors into cells or tissues are available and equally suitable for use *in vivo*, *in vitro*, and *ex vivo*. For *ex vivo* therapy, vectors may be introduced into stem cells taken from the patient and clonally propagated for autologous transplant back into that same patient. Delivery by transfection, by liposome injections or polycationic amino polymers (Goldman, C.K. et al. (1997) Nature Biotechnology 15:462-66; incorporated herein by reference) may be achieved using methods which are well known in the art.

Any of the therapeutic methods described above may be applied to any subject in need of such therapy, including, for example, mammals such as dogs, cats, cows, horses, rabbits, monkeys, and most preferably, humans.

An additional embodiment of the invention relates to the administration of a pharmaceutical composition, in conjunction with a pharmaceutically acceptable carrier, for any of the therapeutic effects discussed above. Such pharmaceutical compositions may consist of HRM, antibodies to HRM, mimetics, agonists, antagonists, or inhibitors of HRM.

The compositions may be administered alone or in combination with at least one other agent, such as stabilizing compound, which may be administered in any sterile, biocompatible pharmaceutical carrier, including, but not limited to, saline, buffered saline, dextrose, and water. The compositions may be administered to a patient alone, or in combination with other agents, drugs or hormones.

The pharmaceutical compositions utilized in this invention may be administered by any number of routes including, but not limited to, oral, intravenous, intramuscular, intra-arterial, intramedullary, intrathecal, intraventricular, transdermal, subcutaneous, intraperitoneal, intranasal, enteral, topical, sublingual, or rectal means.

In addition to the active ingredients, these pharmaceutical compositions may contain suitable pharmaceutically-acceptable carriers comprising excipients and auxiliaries which facilitate processing of the active compounds into preparations which can be used pharmaceutically. Further details on techniques for formulation and administration may be found in the latest edition of Remington's Pharmaceutical Sciences (Maack Publishing Co., Easton PA).

Pharmaceutical compositions for oral administration can be formulated using pharmaceutically acceptable carriers well known in the art in dosages suitable for oral administration. Such carriers enable the pharmaceutical compositions to be formulated as tablets, pills, dragees, capsules, liquids, gels, syrups, slurries, suspensions, and the like, for ingestion by the patient.

Pharmaceutical preparations for oral use can be obtained through combination of active compounds with solid excipient, optionally grinding a resulting mixture, and processing the mixture of granules, after adding suitable auxiliaries, if desired, to obtain tablets or dragee cores. Suitable excipients are carbohydrate or protein fillers, such as sugars, including lactose, sucrose, mannitol, or sorbitol; starch from corn, wheat, rice, potato, or other plants; cellulose, such as methyl cellulose, hydroxypropylmethyl-cellulose, or sodium carboxymethylcellulose; gums including arabic and tragacanth; and proteins such as gelatin and collagen. If desired, disintegrating or solubilizing agents may be added, such as the cross-linked polyvinyl pyrrolidone, agar, alginic acid, or a salt thereof, such as sodium alginate.

Dragee cores may be used in conjunction with suitable coatings, such as concentrated sugar solutions, which may also contain gum arabic, talc, polyvinylpyrrolidone, carbopol gel,

polyethylene glycol, and/or titanium dioxide, lacquer solutions, and suitable organic solvents or solvent mixtures. Dyestuffs or pigments may be added to the tablets or dragee coatings for product identification or to characterize the quantity of active compound, i.e., dosage.

Pharmaceutical preparations which can be used orally include push-fit capsules made of gelatin, as well as soft, sealed capsules made of gelatin and a coating, such as glycerol or sorbitol. Push-fit capsules can contain active ingredients mixed with a filler or binders, such as lactose or starches, lubricants, such as talc or magnesium stearate, and, optionally, stabilizers. In soft capsules, the active compounds may be dissolved or suspended in suitable liquids, such as fatty oils, liquid, or liquid polyethylene glycol with or without stabilizers.

Pharmaceutical formulations suitable for parenteral administration may be formulated in aqueous solutions, preferably in physiologically compatible buffers such as Hanks' solution, Ringer's solution, or physiologically buffered saline. Aqueous injection suspensions may contain substances which increase the viscosity of the suspension, such as sodium carboxymethyl cellulose, sorbitol, or dextran. Additionally, suspensions of the active compounds may be prepared as appropriate oily injection suspensions. Suitable lipophilic solvents or vehicles include fatty oils such as sesame oil, or synthetic fatty acid esters, such as ethyl oleate or triglycerides, or liposomes. Non-lipid polycationic amino polymers may also be used for delivery. Optionally, the suspension may also contain suitable stabilizers or agents which increase the solubility of the compounds to allow for the preparation of highly concentrated solutions.

For topical or nasal administration, penetrants appropriate to the particular barrier to be permeated are used in the formulation. Such penetrants are generally known in the art.

The pharmaceutical compositions of the present invention may be manufactured in a manner that is known in the art, e.g., by means of conventional mixing, dissolving, granulating, dragee-making, levigating, emulsifying, encapsulating, entrapping, or lyophilizing processes.

The pharmaceutical composition may be provided as a salt and can be formed with many acids, including but not limited to, hydrochloric, sulfuric, acetic, lactic, tartaric, malic, succinic, etc. Salts tend to be more soluble in aqueous or other protonic solvents than are the corresponding free base forms. In other cases, the preferred preparation may be a lyophilized powder which may contain any or all of the following: 1-50 mM histidine, 0.1%-2% sucrose, and 2-7% mannitol, at a pH range of 4.5 to 5.5, that is combined with buffer prior to use.

After pharmaceutical compositions have been prepared, they can be placed in an appropriate container and labeled for treatment of an indicated condition. For administration of HRM, such labeling would include amount, frequency, and method of administration.

Pharmaceutical compositions suitable for use in the invention include compositions
5 wherein the active ingredients are contained in an effective amount to achieve the intended purpose. The determination of an effective dose is well within the capability of those skilled in the art.

For any compound, the therapeutically effective dose can be estimated initially either in cell culture assays, e.g., of neoplastic cells, or in animal models, usually mice, rabbits,
10 dogs, or pigs. The animal model may also be used to determine the appropriate concentration range and route of administration. Such information can then be used to determine useful doses and routes for administration in humans.

A therapeutically effective dose refers to that amount of active ingredient, for example HRM or fragments thereof, antibodies of HRM, agonists, antagonists or inhibitors of HRM,
15 which ameliorates the symptoms or condition. Therapeutic efficacy and toxicity may be determined by standard pharmaceutical procedures in cell cultures or experimental animals, e.g., ' (the dose therapeutically effective in 50% of the population) and LD50 (the dose lethal to 50% of the population). The dose ratio between therapeutic and toxic effects is the therapeutic index, and it can be expressed as the ratio, ED50/LD50.

Pharmaceutical compositions which exhibit large therapeutic indices are preferred.
20 The data obtained from cell culture assays and animal studies is used in formulating a range of dosage for human use. The dosage contained in such compositions is preferably within a range of circulating concentrations that include the ED50 with little or no toxicity. The dosage varies within this range depending upon the dosage form employed, sensitivity of the
25 patient, and the route of administration.

The exact dosage will be determined by the practitioner, in light of factors related to the subject that requires treatment. Dosage and administration are adjusted to provide sufficient levels of the active moiety or to maintain the desired effect. Factors which may be taken into account include the severity of the disease state, general health of the subject, age,
30 weight, and gender of the subject, diet, time and frequency of administration, drug combination(s), reaction sensitivities, and tolerance/response to therapy. Long-acting pharmaceutical compositions may be administered every 3 to 4 days, every week, or once

every two weeks depending on half-life and clearance rate of the particular formulation.

Normal dosage amounts may vary from 0.1 to 100,000 micrograms, up to a total dose of about 1 g, depending upon the route of administration. Guidance as to particular dosages and methods of delivery is provided in the literature and generally available to practitioners in the art. Those skilled in the art will employ different formulations for nucleotides than for proteins or their inhibitors. Similarly, delivery of polynucleotides or polypeptides will be specific to particular cells, conditions, locations, etc.

DIAGNOSTICS

In another embodiment, antibodies which specifically bind HRM may be used for the diagnosis of conditions or diseases characterized by expression of HRM, or in assays to monitor patients being treated with HRM, agonists, antagonists or inhibitors. The antibodies useful for diagnostic purposes may be prepared in the same manner as those described above for therapeutics. Diagnostic assays for HRM include methods which utilize the antibody and a label to detect HRM in human body fluids or extracts of cells or tissues. The antibodies may be used with or without modification, and may be labeled by joining them, either covalently or non-covalently, with a reporter molecule. A wide variety of reporter molecules which are known in the art may be used, several of which are described above.

A variety of protocols including ELISA, RIA, and FACS for measuring HRM are known in the art and provide a basis for diagnosing altered or abnormal levels of HRM expression. Normal or standard values for HRM expression are established by combining body fluids or cell extracts taken from normal mammalian subjects, preferably human, with antibody to HRM under conditions suitable for complex formation. The amount of standard complex formation may be quantified by various methods, but preferably by photometric means. Quantities of HRM expressed in subject, control and disease, samples are compared with the standard values. Deviation between standard and subject values establishes the parameters for diagnosing disease.

In another embodiment of the invention, the polynucleotides encoding HRM may be used for diagnostic purposes. The polynucleotides which may be used include oligonucleotide sequences, complementary RNA and DNA molecules, and PNAs. The polynucleotides may be used to detect and quantitate gene expression in biopsied tissues in which expression of HRM may be correlated with disease. The diagnostic assay may be used to distinguish between absence, presence, and excess expression of HRM, and to monitor

regulation of HRM levels during therapeutic intervention.

In one aspect, hybridization with PCR probes which are capable of detecting polynucleotide sequences, including genomic sequences, encoding HRM or closely related molecules, may be used to identify nucleic acid sequences which encode HRM. The
5 specificity of the probe, whether it is made from a highly specific region, e.g., 10 unique nucleotides in the 5' regulatory region, or a less specific region, e.g., especially in the 3' coding region, and the stringency of the hybridization or amplification (maximal, high, intermediate, or low) will determine whether the probe identifies only naturally occurring sequences encoding HRM, alleles, or related sequences.

10 Probes may also be used for the detection of related sequences, and should preferably contain at least 50% of the nucleotides from any of the HRM encoding sequences. The hybridization probes of the subject invention may be DNA or RNA and derived from the nucleotide sequence of SEQ ID Nos: 11-20 or from genomic sequence including promoter, enhancer elements, and introns of the naturally occurring HRM.

15 Means for producing specific hybridization probes for DNAs encoding HRM include the cloning of nucleic acid sequences encoding HRM or HRM derivatives into vectors for the production of mRNA probes. Such vectors are known in the art, commercially available, and may be used to synthesize RNA probes *in vitro* by means of the addition of the appropriate RNA polymerases and the appropriate labeled nucleotides. Hybridization probes may be
20 labeled by a variety of reporter groups, for example, radionuclides such as ^{32}P or ^{35}S , or enzymatic labels, such as alkaline phosphatase coupled to the probe via avidin/biotin coupling systems, and the like.

Polynucleotide sequences encoding HRM may be used for the diagnosis of conditions, disorders, or diseases which are associated with either increased or decreased expression of
25 HRM. Examples of such conditions or diseases include adenocarcinoma, leukemia, lymphoma, melanoma, myeloma, sarcoma, teratocarcinoma, and cancers of the adrenal gland, bladder, bone, brain, breast, cervix, gall bladder, ganglia, gastrointestinal tract, heart, kidney, liver, lung, bone marrow, muscle, ovary, pancreas, parathyroid, penis, prostate, salivary glands, skin, spleen, testis, thymus, thyroid, and uterus; and immune disorders such as AIDS,
30 Addison's disease, adult respiratory distress syndrome, allergies, anemia, asthma, atherosclerosis, bronchitis, cholecystitis, Crohn's disease, ulcerative colitis, atopic dermatitis, dermatomyositis, diabetes mellitus, emphysema, atrophic gastritis, glomerulonephritis, gout,

Graves' disease, hypereosinophilia, irritable bowel syndrome, lupus erythematosus, multiple sclerosis, myasthenia gravis, myocardial or pericardial inflammation, osteoarthritis, osteoporosis, pancreatitis, polymyositis, rheumatoid arthritis, scleroderma, Sjögren's syndrome, and thyroiditis. The polynucleotide sequences encoding HRM may be used in
5 Southern or northern analysis, dot blot, or other membrane-based technologies; in PCR technologies; or in dipstick, pin, ELISA-like assays or microarrays utilizing fluids or tissues from patient biopsies to detect altered HRM expression. Such qualitative or quantitative methods are well known in the art.

In a particular aspect, the nucleotide sequences encoding HRM may be useful in
10 assays that detect activation or induction of various cancers, particularly those mentioned above. The nucleotide sequences encoding HRM may be labeled by standard methods, and added to a fluid or tissue sample from a patient under conditions suitable for the formation of hybridization complexes. After a suitable incubation period, the sample is washed and the signal is quantitated and compared with a standard value. If the amount of signal in the
15 biopsied or extracted sample is significantly altered from that of a comparable control sample, the nucleotide sequences have hybridized with nucleotide sequences in the sample, and the presence of altered levels of nucleotide sequences encoding HRM in the sample indicates the presence of the associated disease. Such assays may also be used to evaluate the efficacy of a particular therapeutic treatment regimen in animal studies, in clinical trials, or in monitoring
20 the treatment of an individual patient.

In order to provide a basis for the diagnosis of disease associated with expression of HRM, a normal or standard profile for expression is established. This may be accomplished by combining body fluids or cell extracts taken from normal subjects, either animal or human, with a sequence, or a fragment thereof, which encodes HRM, under conditions suitable for
25 hybridization or amplification. Standard hybridization may be quantified by comparing the values obtained from normal subjects with those from an experiment where a known amount of a substantially purified polynucleotide is used. Standard values obtained from normal samples may be compared with values obtained from samples from patients who are symptomatic for disease. Deviation between standard and subject values is used to establish
30 the presence of disease.

Once disease is established and a treatment protocol is initiated, hybridization assays may be repeated on a regular basis to evaluate whether the level of expression in the patient

begins to approximate that which is observed in the normal patient. The results obtained from successive assays may be used to show the efficacy of treatment over a period ranging from several days to months.

With respect to cancer, the presence of a relatively high amount of transcript in
5 biopsied tissue from an individual may indicate a predisposition for the development of the disease or may provide a means for detecting the disease prior to the appearance of actual clinical symptoms. A more definitive diagnosis of this type may allow health professionals to employ preventative measures or aggressive treatment earlier thereby preventing the development or further progression of the cancer.

10 Additional diagnostic uses for oligonucleotides designed from the sequences encoding HRM may involve the use of PCR. Such oligomers may be chemically synthesized, generated enzymatically, or produced *in vitro*. Oligomers will preferably consist of two nucleotide sequences, one with sense orientation (5'→3') and another with antisense (3'←5'), employed under optimized conditions for identification of a specific gene or condition. The
15 same two oligomers, nested sets of oligomers, or even a degenerate pool of oligomers may be employed under less stringent conditions for detection and/or quantitation of closely related DNA or RNA sequences.

Methods which may also be used to quantitate the expression of HRM include radiolabeling or biotinylating nucleotides, coamplification of a control nucleic acid, and
20 standard curves onto which the experimental results are interpolated (Melby, P.C. et al. (1993) J. Immunol. Methods, 159:235-244; Duplaa, C. et al. (1993) Anal. Biochem. 229-236). The speed of quantitation of multiple samples may be accelerated by running the assay in an ELISA-like format where the oligomer of interest is presented in various dilutions and a spectrophotometric or colorimetric response gives rapid quantitation.

25 In further embodiments, oligonucleotides or longer fragments derived from any of the polynucleotide sequences described herein may be used as targets in a microarray. The microarray can be used to monitor the expression level of large numbers of genes simultaneously (to produce a transcript image), and to identify genetic variants, mutations and polymorphisms. This information may be used to determine gene function, understanding the
30 genetic basis of disease, diagnosing disease, and in developing and in monitoring the activities of therapeutic agents.

In one embodiment, the microarray is prepared and used according to the methods

described in PCT application WO95/11995 (Chee et al.), Lockhart, D. J. et al. (1996; Nat. Biotech. 14: 1675-1680) and Schena, M. et al. (1996; Proc. Natl. Acad. Sci. 93: 10614-10619), all of which are incorporated herein in their entirety by reference.

The microarray is preferably composed of a large number of unique, single-stranded
5 nucleic acid sequences, usually either synthetic antisense oligonucleotides or fragments of
cDNAs, fixed to a solid support. The oligonucleotides are preferably about 6-60 nucleotides
in length, more preferably 15-30 nucleotides in length, and most preferably about 20-25
nucleotides in length. For a certain type of microarray, it may be preferable to use
oligonucleotides which are only 7-10 nucleotides in length. The microarray may contain
10 oligonucleotides which cover the known 5', or 3', sequence, or contain sequential
oligonucleotides which cover the full length sequence; or unique oligonucleotides selected
from particular areas along the length of the sequence. Polynucleotides used in the
microarray may be oligonucleotides that are specific to a gene or genes of interest in which at
least a fragment of the sequence is known or that are specific to one or more unidentified
15 cDNAs which are common to a particular cell or tissue type or to a normal, developmental, or
disease state. In certain situations it may be appropriate to use pairs of oligonucleotides on a
microarray. The "pairs" will be identical, except for one nucleotide which preferably is
located in the center of the sequence. The second oligonucleotide in the pair (mismatched by
one) serves as a control. The number of oligonucleotide pairs may range from 2 to one
20 million.

In order to produce oligonucleotides to a known sequence for a microarray, the gene
of interest is examined using a computer algorithm which starts at the 5' or more preferably at
the 3' end of the nucleotide sequence. The algorithm identifies oligomers of defined length
that are unique to the gene, have a GC content within a range suitable for hybridization, and
25 lack predicted secondary structure that may interfere with hybridization. In one aspect, the
oligomers are synthesized at designated areas on a substrate using a light-directed chemical
process. The substrate may be paper, nylon or other type of membrane, filter, chip, glass
slide, or any other suitable solid support.

In another aspect, the oligonucleotides may be synthesized on the surface of the
30 substrate by using a chemical coupling procedure and an ink jet application apparatus, as
described in PCT application WO95/251116 (Baldeschweiler et al.) which is incorporated
herein in its entirety by reference. In another aspect, a "gridded" array analogous to a dot or

slot blot (HYBRIDOT apparatus, Life Technologies) may be used to arrange and link cDNA fragments or oligonucleotides to the surface of a substrate using a vacuum system, thermal, UV, mechanical or chemical bonding procedures. In yet another aspect, an array may be produced by hand or using available devices, materials, and machines (including Brinkmann
5 multichannel pipettors or robotic instruments) and may contain 8, 24, 96, 384, 1536 or 6144 oligonucleotides, or any other multiple from 2 to one million which lends itself to the efficient use of commercially available instrumentation.

In order to conduct sample analysis using the microarrays, polynucleotides are extracted from a biological sample. The biological samples may be obtained from any bodily
10 fluid (blood, urine, saliva, phlegm, gastric juices, etc.), cultured cells, biopsies, or other tissue preparations. To produce probes, the polynucleotides extracted from the sample are used to produce nucleic acid sequences which are complementary to the nucleic acids on the microarray. If the microarray consists of cDNAs, antisense RNAs (aRNA) are appropriate probes. Therefore, in one aspect, mRNA is used to produce cDNA which, in turn and in the
15 presence of fluorescent nucleotides, is used to produce fragment or oligonucleotide aRNA probes. These fluorescently labeled probes are incubated with the microarray so that the probe sequences hybridize to the cDNA oligonucleotides of the microarray. In another aspect, complementary nucleic acid sequences are used as probes and can also include polynucleotides, fragments, complementary, or antisense sequences produced using
20 restriction enzymes, PCR technologies, and Oligolabelling or TransProbe kits (Amersham Pharmacia) well known in the area of hybridization technology.

Incubation conditions are adjusted so that hybridization occurs with precise complementary matches or with various degrees of less complementarity. After removal of nonhybridized probes, a scanner is used to determine the levels and patterns of fluorescence.
25 The scanned images are examined to determine degree of complementarity and the relative abundance of each oligonucleotide sequence on the microarray. A detection system may be used to measure the absence, presence, and amount of hybridization for all of the distinct sequences simultaneously. This data may be used for large scale correlation studies or functional analysis of the sequences, mutations, variants, or polymorphisms among samples
30 (Heller, R.A. et al., (1997) Proc. Natl. Acad. Sci. 94:2150-55).

In another embodiment of the invention, the nucleic acid sequences which encode HRM may also be used to generate hybridization probes which are useful for mapping the

naturally occurring genomic sequence. The sequences may be mapped to a particular chromosome, to a specific region of a chromosome or to artificial chromosome constructions, such as human artificial chromosomes (HACs), yeast artificial chromosomes (YACs), bacterial artificial chromosomes (BACs), bacterial P1 constructions or single chromosome cDNA libraries as reviewed in Price, C.M. (1993; Blood Rev. 7:127-134) and Trask, B.J. (1991; Trends Genet. 7:149-154).

Fluorescent in situ hybridization (FISH as described in Verma et al. (1988) Human Chromosomes: A Manual of Basic Techniques, Pergamon Press, New York NY) may be correlated with other physical chromosome mapping techniques and genetic map data. Examples of genetic map data can be found in various scientific journals or at Online Mendelian Inheritance in Man (OMIM). Correlation between the location of the gene encoding HRM on a physical chromosomal map and a specific disease, or predisposition to a specific disease, may help delimit the region of DNA associated with that genetic disease. The nucleotide sequences of the subject invention may be used to detect differences in gene sequences between normal, carrier, or affected individuals.

In situ hybridization of chromosomal preparations and physical mapping techniques such as linkage analysis using established chromosomal markers may be used for extending genetic maps. Often the placement of a gene on the chromosome of another mammalian species, such as mouse, may reveal associated markers even if the number or arm of a particular human chromosome is not known. New sequences can be assigned to chromosomal arms, or parts thereof, by physical mapping. This provides valuable information to investigators searching for disease genes using positional cloning or other gene discovery techniques. Once the disease or syndrome has been crudely localized by genetic linkage to a particular genomic region, for example, AT to 11q22-23 (Gatti, R.A. et al. (1988) Nature 336:577-580), any sequences mapping to that area may represent associated or regulatory genes for further investigation. The nucleotide sequence of the subject invention may also be used to detect differences in the chromosomal location due to translocation, inversion, etc. among normal, carrier, or affected individuals.

In another embodiment of the invention, HRM, its catalytic or immunogenic fragments or oligopeptides thereof, can be used for screening libraries of compounds in any of a variety of drug screening techniques. The fragment employed in such screening may be free in solution, affixed to a solid support, borne on a cell surface, or located intracellularly. The

formation of binding complexes, between HRM and the agent being tested, may be measured.

Another technique for drug screening which may be used provides for high throughput screening of compounds having suitable binding affinity to the protein of interest as described in published PCT application WO84/03564. In this method, large numbers of different small
5 test compounds are synthesized on a solid substrate, such as plastic pins or some other surface. The test compounds are reacted with HRM, or fragments thereof, and washed. Bound HRM is then detected by methods well known in the art. Purified HRM can also be coated directly onto plates for use in the aforementioned drug screening techniques. Alternatively, non-neutralizing antibodies can be used to capture the peptide and immobilize
10 it on a solid support.

In another embodiment, one may use competitive drug screening assays in which neutralizing antibodies capable of binding HRM specifically compete with a test compound for binding HRM. In this manner, the antibodies can be used to detect the presence of any peptide which shares one or more antigenic determinants with HRM.

15 In additional embodiments, the nucleotide sequences which encode HRM may be used in any molecular biology techniques that have yet to be developed, provided the new techniques rely on properties of nucleotide sequences that are currently known, including, but not limited to, such properties as the triplet genetic code and specific base pair interactions.

The examples below are provided to illustrate the subject invention and are not
20 included for the purpose of limiting the invention.

EXAMPLES

For purposes of example, the preparation and sequencing of the LNODNOT03 cDNA library, from which Incyte Clones were isolated, is described. Preparation and sequencing of cDNAs in libraries in the LIFESEQ database (Incyte Pharmaceuticals, Palo Alto CA) have
25 varied over time, and the gradual changes involved use of kits, plasmids, and machinery available at the particular time the library was made and analyzed.

I LNODNOT03 cDNA Library Construction

The LNODNOT03 cDNA library was constructed from microscopically normal lymph node tissue excised from a 67-year-old Caucasian male with lung cancer during a
30 segmental lung resection. This nontumorous tissue was associated with an invasive grade 3 squamous cell carcinoma of the lung. The patient history included hemangioma and tobacco use. The patient was taking DOXYCYCLINE, a tetracycline, to treat an infection.

The frozen tissue was homogenized and lysed using a Brinkmann Homogenizer Polytron PT-3000 (Brinkmann Instruments, Westbury NJ) in guanidinium isothiocyanate solution. The lysate was centrifuged over a 5.7 M CsCl cushion using a Beckman SW28 rotor in a Beckman L8-70M Ultracentrifuge (Beckman Coulter Instruments, Fullerton CA) for 18 hours at 25,000 rpm at ambient temperature. The RNA was extracted with acid phenol pH 4.7, precipitated using 0.3 M sodium acetate and 2.5 volumes of ethanol, resuspended in RNase-free water, and treated with DNase at 37°C. Extraction and precipitation were repeated as before. The mRNA was isolated with the OLIGOTEX kit (QIAGEN, Inc., Chatsworth CA) and used to construct the cDNA library.

The mRNA was handled according to the recommended protocols in the SUPERScript Plasmid System for cDNA synthesis and plasmid cloning (Life Technologies). The cDNAs were fractionated on a SEPHAROSE CL4B column (Amersham Pharmacia), and those cDNAs exceeding 400 bp were ligated into pINCY 1. The plasmid pINCY (Incyte Pharmaceuticals) was subsequently transformed into DH5 α competent cells (Life Technologies).

II Isolation and Sequencing of cDNA Clones

Plasmid DNA was released from the cells and purified using the REAL Prep 96 plasmid kit (QIAGEN, Inc.). This kit enabled the simultaneous purification of 96 samples in a 96-well block using multi-channel reagent dispensers. The recommended protocol was employed except for the following changes: 1) the bacteria were cultured in 1 ml of sterile Terrific Broth (Life Technologies) with carbenicillin at 25 mg/L and glycerol at 0.4%; 2) after inoculation, the cultures were incubated for 19 hours and at the end of incubation, the cells were lysed with 0.3 ml of lysis buffer; and 3) following isopropanol precipitation, the plasmid DNA pellet was resuspended in 0.1 ml of distilled water. After the last step in the protocol, samples were transferred to a 96-well block for storage at 4°C.

The cDNAs were sequenced by the method of Sanger et al. (1975, J. Mol. Biol. 94:441f), using a Micro Lab 2200 (Hamilton, Reno NV) in combination with PTC200 from MJ Research. and ABI 377 DNA sequencing systems (Perkin Elmer); and the reading frame was determined.

III Homology Searching of cDNA Clones and Their Deduced Proteins

The nucleotide sequences and/or amino acid sequences of the Sequence Listing were used to query sequences in the GenBank, SwissProt, BLOCKS, and Pima II databases. These

databases, which contain previously identified and annotated sequences, were searched for regions of homology using BLAST, which stands for Basic Local Alignment Search Tool (Altschul, S.F. (1993) J. Mol. Evol 36:290-300; Altschul, et al. (1990) J. Mol. Biol. 215:403-410).

5 BLAST produced alignments of both nucleotide and amino acid sequences to determine sequence similarity. Because of the local nature of the alignments, BLAST was especially useful in determining exact matches or in identifying homologs which may be of prokaryotic (bacterial) or eukaryotic (animal, fungal, or plant) origin. Other algorithms such as the one described in Smith, T. et al. (1992, Protein Engineering 5:35-51), incorporated
10 herein by reference, could have been used when dealing with primary sequence patterns and secondary structure gap penalties. The sequences disclosed in this application have lengths of at least 49 nucleotides, and no more than 12% uncalled bases (where N is recorded rather than A, C, G, or T).

The BLAST approach searched for matches between a query sequence and a database
15 sequence. BLAST evaluated the statistical significance of any matches found, and reported only those matches that satisfy the user-selected threshold of significance. In this application, threshold was set at 10^{-25} for nucleotides and 10^{-14} for peptides.

Incyte nucleotide sequences were searched against the GenBank databases for primate (pri), rodent (rod), and other mammalian sequences (mam); and deduced amino acid
20 sequences from the same clones were then searched against GenBank functional protein databases, mammalian (mamp), vertebrate (vrtp), and eukaryote (eukp) for homology.

IV Northern Analysis

Northern analysis is a laboratory technique used to detect the presence of a transcript of a gene and involves the hybridization of a labeled nucleotide sequence to a membrane on
25 which RNAs from a particular cell type or tissue have been bound (Sambrook et al., supra).

Analogous computer techniques use BLAST to search for identical or related molecules in nucleotide databases such as GenBank or the LIFESEQ database (Incyte Pharmaceuticals). This analysis is much faster than multiple, membrane-based hybridizations. In addition, the sensitivity of the computer search can be modified to
30 determine whether any particular match is categorized as exact or homologous.

The basis of the search is the product score which is defined as:

$$\frac{\% \text{ sequence identity} \times \% \text{ maximum BLAST score}}{100}$$

The product score takes into account both the degree of similarity between two sequences and the length of the sequence match. For example, with a product score of 40, the match will be exact within a 1-2% error; and at 70, the match will be exact. Homologous molecules are usually identified by selecting those which show product scores between 15 and 40, although
5 lower scores may identify related molecules.

The results of northern analysis are reported as a list of libraries in which the transcript encoding HRM occurs. Abundance and percent abundance are also reported. Abundance directly reflects the number of times a particular transcript is represented in a cDNA library, and percent abundance is abundance divided by the total number of sequences
10 examined in the cDNA library.

V Extension of HRM Encoding Polynucleotides

The nucleic acid sequence of an Incyte Clone disclosed in the Sequence Listing was used to design oligonucleotide primers for extending a partial nucleotide sequence to full length. One primer was synthesized to initiate extension in the antisense direction, and the
15 other was synthesized to extend sequence in the sense direction. Primers were used to facilitate the extension of the known sequence "outward" generating amplicons containing new, unknown nucleotide sequence for the region of interest. The initial primers were designed from the cDNA using OLIGO 4.06 (National Biosciences), or another appropriate program, to be about 22 to about 30 nucleotides in length, to have a GC content of 50% or
20 more, and to anneal to the target sequence at temperatures of about 68° to about 72°C. Any stretch of nucleotides which would result in hairpin structures and primer-primer dimerizations was avoided.

Selected human cDNA libraries (Life Technologies) were used to extend the sequence. If more than one extension is necessary or desired, additional sets of primers are designed to
25 further extend the known region.

High fidelity amplification was obtained by following the instructions for the XL-PCR kit (Perkin Elmer) and thoroughly mixing the enzyme and reaction mix. Beginning with 40 pmol of each primer and the recommended concentrations of all other components of the kit, PCR was performed using the PTC200 (MJ Research) and the following parameters: Step 1,
30 94° C for 1 min (initial denaturation); Step 2, 65° C for 1 min; Step 3, 68° C for 6 min; Step 4, 94° C for 15 sec; Step 5, 65° C for 1 min; Step 6, 68° C for 7 min; Step 7, for 15 additional cycles; Step 8, 94° C for 15 sec; Step 9, 65° C for 1 min; Step 10, 68° C for 7:15 min; Step

11, Repeat step 8-10 for 12 cycles; Step 12, 72° C for 8 min; and Step 13, hold at 4° C.

A 5-10 μ l aliquot of the reaction mixture was analyzed by electrophoresis on a low concentration (about 0.6-0.8%) agarose mini-gel to determine which reactions were successful in extending the sequence. Bands thought to contain the largest products were excised from the gel, purified using QIAQUICK kit (QIAGEN Inc.), and trimmed of overhangs using Klenow enzyme to facilitate religation and cloning.

After ethanol precipitation, the products were redissolved in 13 μ l of ligation buffer, 1 μ l T4-DNA ligase (15 units) and 1 μ l T4 polynucleotide kinase were added, and the mixture was incubated at room temperature for 2-3 hours or overnight at 16° C. Competent *E. coli* cells (in 40 μ l of appropriate media) were transformed with 3 μ l of ligation mixture and cultured in 80 μ l of SOC medium (Sambrook et al., supra). After incubation for one hour at 37° C, the *E. coli* mixture was plated on Luria Bertani (LB)-agar (Sambrook et al., supra) containing 2x Carb. The following day, several colonies were randomly picked from each plate and cultured in 150 μ l of liquid LB/2x Carb medium placed in an individual well of an appropriate, commercially-available, sterile 96-well microtiter plate. The following day, 5 μ l of each overnight culture was transferred into a non-sterile 96-well plate and after dilution 1:10 with water, 5 μ l of each sample was transferred into a PCR array.

For PCR amplification, 18 μ l of concentrated PCR reaction mix (3.3x) containing 4 units of rTth DNA polymerase, a vector primer, and one or both of the gene specific primers used for the extension reaction were added to each well. Amplification was performed using the following conditions: Step 1, 94° C for 60 sec; Step 2, 94° C for 20 sec; Step 3, 55° C for 30 sec; Step 4, 72° C for 90 sec; Step 5, Repeat steps 2-4 for an additional 29 cycles; Step 6, 72° C for 180 sec; and Step 7, hold at 4° C.

Aliquots of the PCR reactions were run on agarose gels together with molecular weight markers. The sizes of the PCR products were compared to the original partial cDNAs, and appropriate clones were selected, ligated into plasmid, and sequenced.

In like manner, the nucleotide sequence of SEQ ID NOs:11-20 are used to obtain 5' regulatory sequences using the procedure above, oligonucleotides designed for 5' extension, and an appropriate genomic library.

30 VI Labeling and Use of Individual Hybridization Probes

Hybridization probes derived from SEQ ID NOs:11-20 are employed to screen cDNAs, genomic DNAs, or mRNAs. Although the labeling of oligonucleotides, consisting of

about 20 base-pairs, is specifically described, essentially the same procedure is used with larger nucleotide fragments. Oligonucleotides are designed using state-of-the-art software such as OLIGO 4.06 (National Biosciences), labeled by combining 50 pmol of each oligomer and 250 μ Ci of [γ - 32 P] adenosine triphosphate (Amersham Pharmacia) and T4 polynucleotide kinase (DuPont NEN, Boston MA). The labeled oligonucleotides are substantially purified with SEPHADEX G-25 superfine resin column (Amersham Pharmacia). A aliquot containing 10^7 counts per minute of the labeled probe is used in a typical membrane-based hybridization analysis of human genomic DNA digested with one of the following endonucleases (Ase I, Bgl II, Eco RI, Pst I, Xba I, or Pvu II; DuPont NEN).

The DNA from each digest is fractionated on a 0.7 percent agarose gel and transferred to nylon membranes (Nytran Plus, Schleicher & Schuell, Durham NH). Hybridization is carried out for 16 hours at 40°C. To remove nonspecific signals, blots are sequentially washed at room temperature under increasingly stringent conditions up to 0.1 x saline sodium citrate and 0.5% sodium dodecyl sulfate. After XOMAT AR film (Eastman Kodak, Rochester NY) is exposed to the blots in a Phosphorimager cassette (Molecular Dynamics, Sunnyvale CA) for several hours, hybridization patterns are compared visually.

VII Microarrays

To produce oligonucleotides for a microarray, SEQ ID NOs:11-20 are examined using a computer algorithm which starts at the 3' end of the nucleotide sequence. The algorithm identified oligomers of defined length that are unique to the gene, have a GC content within a range suitable for hybridization, and lack predicted secondary structure that would interfere with hybridization. The algorithm identifies approximately 20 sequence-specific oligonucleotides of 20 nucleotides in length (20-mers). A matched set of oligonucleotides are created in which one nucleotide in the center of each sequence is altered. This process is repeated for each gene in the microarray, and double sets of twenty 20 mers are synthesized and arranged on the surface of the silicon chip using a light-directed chemical process (Chee, M. et al., PCT/WO95/11995, incorporated herein by reference).

In the alternative, a chemical coupling procedure and an ink jet device are used to synthesize oligomers on the surface of a substrate (Baldeschweiler, J.D. et al., PCT/WO95/25116, incorporated herein by reference). In another alternative, a "gridded" array analogous to a dot (or slot) blot is used to arrange and link cDNA fragments or oligonucleotides to the surface of a substrate using a vacuum system, thermal, UV,

mechanical or chemical bonding procedures. A typical array may be produced by hand or using available materials and machines and contain grids of 8 dots, 24 dots, 96 dots, 384 dots, 1536 dots or 6144 dots. After hybridization, the microarray is washed to remove nonhybridized probes, and a scanner is used to determine the levels and patterns of fluorescence. The scanned image is examined to determine degree of complementarity and the relative abundance/expression level of each oligonucleotide sequence in the microarray.

VIII Complementary Polynucleotides

Sequence complementary to the sequence encoding HRM, or any part thereof, is used to detect, decrease or inhibit expression of naturally occurring HRM. Although use of oligonucleotides comprising from about 15 to about 30 base-pairs is described, essentially the same procedure is used with smaller or larger sequence fragments. Appropriate oligonucleotides are designed using OLIGO 4.06 software and the coding sequence of HRM, SEQ ID NOs:11-20. To inhibit transcription, a complementary oligonucleotide is designed from the most unique 5' sequence and used to prevent promoter binding to the coding sequence. To inhibit translation, a complementary oligonucleotide is designed to prevent ribosomal binding to the transcript encoding HRM.

IX Expression of HRM

Expression of HRM is accomplished by subcloning the cDNAs into appropriate vectors and transforming the vectors into host cells. In this case, the cloning vector is also used to express HRM in *E. coli*. Upstream of the cloning site, this vector contains a promoter for β -galactosidase, followed by sequence containing the amino-terminal Met, and the subsequent seven residues of β -galactosidase. Immediately following these eight residues is a bacteriophage promoter useful for transcription and a linker containing a number of unique restriction sites.

Induction of an isolated, transformed bacterial strain with IPTG using standard methods produces a fusion protein which consists of the first eight residues of β -galactosidase, about 5 to 15 residues of linker, and the full length protein. The signal residues direct the secretion of HRM into the bacterial growth media which can be used directly in the following assay for activity.

X Demonstration of HRM Activity

HRM can be expressed in a mammalian cell line such as DLD-1 or HCT116 (ATCC) by transforming the cells with a eukaryotic expression vector encoding HRM. Eukaryotic

expression vectors are commercially available and the techniques to introduce them into cells are well known to those skilled in the art. The effect of HRM on cell morphology may be visualized by microscopy; the effect on cell growth may be determined by measuring cell doubling-time; and the effect on tumorigenicity may be assessed by the ability of transformed
5 cells to grow in a soft agar growth assay (Grodén, J. et al. (1995) Cancer Res. 55:1531-1539).

XI Production of HRM Specific Antibodies

HRM that is substantially purified using PAGE electrophoresis (Sambrook, supra), or other purification techniques, is used to immunize rabbits and to produce antibodies using standard protocols. The amino acid sequence deduced from SEQ ID NOs:11-20 is analyzed
10 using LASERGENE software (DNASTAR Inc) to determine regions of high immunogenicity and a corresponding oligopeptide is synthesized and used to raise antibodies by means known to those of skill in the art. Selection of appropriate epitopes, such as those near the C-terminus or in hydrophilic regions, is described by Ausubel et al. (supra), and others.

Typically, the oligopeptides are 15 residues in length, synthesized using an Applied
15 Biosystems Peptide Synthesizer Model 431A using fmoc-chemistry, and coupled to keyhole limpet hemocyanin (KLH, Sigma-Aldrich, St. Louis MO) by reaction with N-maleimidobenzoyl-N-hydroxysuccinimide ester (MBS; Ausubel et al., supra). Rabbits are immunized with the oligopeptide-KLH complex in complete Freund's adjuvant. The resulting antisera are tested for anti-peptide activity, for example, by binding the peptide to plastic,
20 blocking with 1% BSA, reacting with rabbit antisera, washing, and reacting with radioiodinated, goat anti-rabbit IgG.

XII Purification of Naturally Occurring HRM Using Specific Antibodies

Naturally occurring or recombinant HRM is substantially purified by immunoaffinity chromatography using antibodies specific for HRM. An immunoaffinity column is
25 constructed by covalently coupling HRM antibody to an activated chromatographic resin, such as CNBr-activated SEPHAROSE (Amersham Pharmacia). After the coupling, the resin is blocked and washed according to the manufacturer's instructions.

Media containing HRM is passed over the immunoaffinity column, and the column is washed under conditions that allow the preferential absorbance of HRM (e.g., high ionic
30 strength buffers in the presence of detergent). The column is eluted under conditions that disrupt antibody/protein binding (eg, a buffer of pH 2-3 or a high concentration of a chaotrope, such as urea or thiocyanate ion), and HRM is collected.

XIII Identification of Molecules Which Interact with HRM

HRM or biologically active fragments thereof are labeled with ^{125}I Bolton-Hunter reagent (Bolton et al. (1973) Biochem. J. 133: 529). Candidate molecules previously arrayed in the wells of a multi-well plate are incubated with the labeled HRM, washed and any wells
5 with labeled HRM complex are assayed. Data obtained using different concentrations of HRM are used to calculate values for the number, affinity, and association of HRM with the candidate molecules.

All publications and patents mentioned in the above specification are herein incorporated by reference. Various modifications and variations of the described method and
10 system of the invention will be apparent to those skilled in the art without departing from the scope and spirit of the invention. Although the invention has been described in connection with specific preferred embodiments, it should be understood that the invention as claimed should not be unduly limited to such specific embodiments. Indeed, various modifications of the described modes for carrying out the invention which are obvious to those skilled in
15 molecular biology or related fields are intended to be within the scope of the following claims.

What is claimed is:

1. A substantially purified human regulatory molecule (HRM) comprising a polypeptide having an amino acid sequence selected from the group consisting of: SEQ ID NOs:1-10.
- 5 2. An isolated and purified polynucleotide sequence having a nucleic acid sequence encoding the human regulatory molecule of claim 1.
3. An isolated and purified polynucleotide sequence which hybridizes under stringent conditions to the polynucleotide sequence of claim 2.
4. A composition comprising the polynucleotide sequence of claim 2.
- 10 5. An isolated and purified polynucleotide sequence having a nucleic acid sequence selected from the group consisting of: SEQ ID NOs:11-20.
6. A microarray containing at least a fragment of at least one of the polynucleotides encoding a polypeptide having an amino acid sequence having the sequence selected from the group consisting of SEQ ID NOs:1-10.
- 15 7. A microarray containing at least a fragment of at least one of the polynucleotides having a nucleic acid selected from the group consisting of:SEQ ID NOs:11-20.
8. An isolated and purified polynucleotide having a nucleic acid sequence which is complementary to the nucleic acid sequence of the polynucleotide of claim 5.
- 20 9. A composition comprising the polynucleotide of claim 5.
10. An expression vector containing the polynucleotide of claim 2.
11. A host cell containing the vector of claim 10.
12. A method for producing a polypeptide encoding a human regulatory molecule, the method comprising the steps of:
 - 25 a) culturing the host cell of claim 11 under conditions suitable for the expression of the polypeptide; and
 - b) recovering the polypeptide from the host cell culture.
13. A pharmaceutical composition comprising a substantially purified human regulatory molecule of claim 1 in conjunction with a suitable pharmaceutical carrier.
- 30 14. A purified antibody which binds specifically to the human regulatory molecule of claim 1.
15. A purified agonist which modulates the activity of the human regulatory

molecule of claim 1.

16. A purified antagonist which decreases the effect of the human regulatory molecule of claim 1.

17. A method for stimulating cell proliferation comprising administering to a cell
5 an effective amount of the human regulatory molecule of claim 1.

18. A method for preventing a cancer comprising administering to a subject in need of such treatment an effective amount of the pharmaceutical composition of claim 13.

19. A method for treating or preventing a cancer comprising administering to a subject in need of such treatment an effective amount of the antagonist of claim 16.

10 20. A method for treating or preventing an immune response comprising administering to a subject in need of such treatment an effective amount of the antagonist of claim 16.

21. A method for detecting a polynucleotide encoding a human regulatory molecule in a biological sample, the method comprising the steps of:

15 a) hybridizing the polynucleotide of claim 8 to nucleic acid material of a biological sample, thereby forming a hybridization complex; and
b) detecting the hybridization complex, wherein the presence of the hybridization complex correlates with the presence of a polynucleotide encoding a human regulatory molecule in the biological sample.

20 22. A method for detecting expression levels of polynucleotides encoding a human regulatory molecule in a biological sample, the method comprising the steps of:

a) hybridizing the microarray of claim 6 to labeled complementary polynucleotides derived from the biological sample, thereby forming hybridization complexes; and

25 b) determining expression of polynucleotides encoding a human regulatory molecule in the biological sample by identifying the presence of the hybridization complexes.

23. A method for detecting expression levels of polynucleotides encoding a human regulatory molecule in a biological sample, the method comprising the steps of:

30 a) hybridizing the microarray of claim 7 to labeled complementary polynucleotides derived from the biological sample, thereby forming hybridization complexes; and

b) determining expression of polynucleotides encoding a human regulatory molecule in the biological sample by identifying the presence of the hybridization complexes.

24. The method of claim 22, wherein before hybridization, the nucleic acid
5 material of the biological sample is amplified and labeled by the polymerase chain reaction.

<110> INCYTE PHARMACEUTICALS, INC.

LAL, Preeti
HILLMAN, Jennifer L.
BANDMAN, Olga
SHAH, Purvi
AU-YOUNG, Janice
YUE, Henry
GUEGLER, Karl J.

<120> HUMAN REGULATORY MOLECULES

<130> PF-0356 PCT

<140> To Be Assigned

<141> Herewith

<150> US 08/933,750

<151> 1997-09-23

<160> 20

<170> PERL Program

<210> 1

<211> 456

<212> PRT

<213> Homo sapiens

<220> -

<223> 009476

<400> 1

Met	Lys	Ile	Glu	Glu	Val	Lys	Ser	Thr	Thr	Lys	Thr	Gln	Arg	Ile
1				5					10					15
Ala	Ser	His	Ser	His	Val	Lys	Gly	Leu	Gly	Leu	Asp	Glu	Ser	Gly
				20					25					30
Leu	Ala	Lys	Gln	Ala	Ala	Ser	Gly	Leu	Val	Gly	Gln	Glu	Asn	Ala
				35					40					45
Arg	Glu	Ala	Cys	Gly	Val	Ile	Val	Glu	Leu	Ile	Glu	Ser	Lys	Lys
				50					55					60
Met	Ala	Gly	Arg	Ala	Val	Leu	Leu	Ala	Gly	Pro	Pro	Gly	Thr	Gly
				65					70					75
Lys	Thr	Ala	Leu	Ala	Leu	Ala	Ile	Ala	Gln	Glu	Leu	Gly	Ser	Lys
				80					85					90
Val	Pro	Phe	Cys	Pro	Met	Val	Gly	Ser	Glu	Val	Tyr	Ser	Thr	Glu
				95					100					105
Ile	Lys	Lys	Thr	Glu	Val	Leu	Met	Glu	Asn	Phe	Arg	Arg	Ala	Ile
				110					115					120
Gly	Leu	Arg	Ile	Lys	Glu	Thr	Lys	Glu	Val	Tyr	Glu	Gly	Glu	Val
				125					130					135
Thr	Glu	Leu	Thr	Pro	Cys	Glu	Thr	Glu	Asn	Pro	Met	Gly	Gly	Tyr
				140					145					150
Gly	Lys	Thr	Ile	Ser	His	Val	Ile	Ile	Gly	Leu	Lys	Thr	Ala	Lys
				155					160					165
Gly	Thr	Lys	Gln	Leu	Lys	Leu	Asp	Pro	Ser	Ile	Phe	Glu	Ser	Leu
				170					175					180
Gln	Lys	Glu	Arg	Val	Glu	Ala	Gly	Asp	Val	Ile	Tyr	Ile	Glu	Ala
				185					190					195
Asn	Ser	Gly	Ala	Val	Lys	Arg	Gln	Gly	Arg	Cys	Asp	Thr	Tyr	Ala
				200					205					210
Thr	Glu	Phe	Asp	Leu	Glu	Ala	Glu	Glu	Tyr	Val	Pro	Leu	Pro	Lys

Gly Asp Val His	215	Lys Lys Lys Glu Ile	220	Ile Gln Asp Val Thr	225
His Asp Leu Asp	230	Val Ala Asn Ala Arg	235	Gln Gly Gly Gln Asp	240
Ile Leu Ser Met	245	Met Gly Gln Leu Met	250	Pro Lys Lys Thr Glu	255
Ile Thr Asp Lys	260	Leu Arg Gly Glu Ile	265	Asn Lys Val Val Asn Lys	270
Tyr Ile Asp Gln	275	Gly Ile Ala Glu Leu	280	Val Pro Gly Val Leu Phe	285
Val Asp Glu Val	290	His Met Leu Asp Ile	295	Glu Cys Phe Thr Tyr Leu	300
His Arg Ala Leu	305	Glu Ser Ser Ile Ala	310	Pro Ile Val Ile Phe Ala	315
Ser Asn Arg Gly	320	Asn Cys Val Ile Arg	325	Gly Thr Glu Asp Ile Thr	330
Ser Pro His Gly	335	Ile Pro Leu Asp Leu	340	Leu Asp Arg Val Met Ile	345
Ile Arg Thr Met	350	Leu Tyr Thr Pro Gln	355	Glu Met Lys Gln Ile Ile	360
Lys Ile Arg Ala	365	Gln Thr Glu Gly Ile	370	Asn Ile Ser Glu Glu Ala	375
Leu Asn His Leu	380	Gly Glu Ile Gly Thr	385	Lys Thr Thr Leu Arg Tyr	390
Ser Val Gln Leu	395	Leu Thr Pro Ala Asn	400	Leu Leu Ala Lys Ile Asn	405
Gly Lys Asp Ser	410	Ile Glu Lys Glu His	415	Val Glu Glu Ile Ser Glu	420
Leu Phe Tyr Asp	425	Ala Lys Ser Ser Ala	430	Lys Ile Leu Ala Asp Gln	435
Gln Asp Lys Tyr	440	Met Lys	445		450
	455				

<210> 2
 <211> 188
 <212> PRT
 <213> Homo sapiens

<220> -
 <223> 077180

<400> 2
 Met Ala Leu Ala Met Leu Val Leu Val Val Ser Pro Trp Ser Ala
 1 5 10 15
 Ala Arg Gly Val Leu Arg Asn Tyr Trp Glu Arg Leu Leu Arg Lys
 20 25 30
 Leu Pro Gln Ser Arg Pro Gly Phe Pro Ser Pro Pro Trp Gly Pro
 35 40 45
 Ala Leu Ala Val Gln Gly Pro Ala Met Phe Thr Glu Pro Ala Asn
 50 55 60
 Asp Thr Ser Gly Ser Lys Glu Asn Ser Ser Leu Leu Asp Ser Ile
 65 70 75
 Phe Trp Met Ala Ala Pro Lys Asn Arg Arg Thr Ile Glu Val Asn
 80 85 90
 Arg Cys Arg Arg Arg Asn Pro Gln Lys Leu Ile Lys Val Lys Asn
 95 100 105
 Asn Ile Asp Val Cys Pro Glu Cys Gly His Leu Lys Gln Lys His
 110 115 120
 Val Leu Cys Ala Tyr Cys Tyr Glu Lys Val Cys Lys Glu Thr Ala
 125 130 135
 Glu Ile Arg Arg Gln Ile Gly Lys Gln Glu Gly Gly Pro Phe Lys

	140		145		150
Ala Pro Thr Ile	Glu Thr Val Val Leu	Tyr Thr Gly Glu Thr	Pro		
	155		160		165
Ser Glu Gln Asp	Gln Gly Lys Arg Ile	Ile Glu Arg Asp Arg	Lys		
	170		175		180
Arg Pro Ser Trp	Phe Thr Gln Asn				
	185				

<210> 3
 <211> 348
 <212> PRT
 <213> Homo sapiens

<220> -
 <223> 118160CD1

<400> 3
 Met Gly Gln Glu Glu Glu Leu Leu Arg Ile Ala Lys Lys Leu Glu
 1 5 10 15
 Lys Met Val Ala Arg Lys Asn Thr Glu Gly Ala Leu Asp Leu Leu
 20 25 30
 Lys Lys Leu His Ser Cys Gln Met Ser Ile Gln Leu Leu Gln Thr
 35 40 45
 Thr Arg Ile Gly Val Ala Val Asn Gly Val Arg Lys His Cys Ser
 50 55 60
 Asp Lys Glu Val Val Ser Leu Ala Lys Val Leu Ile Lys Asn Trp
 65 70 75
 Lys Arg Leu Leu Asp Ser Pro Gly Pro Pro Lys Gly Glu Lys Gly
 80 85 90
 Glu Glu Arg Glu Lys Ala Lys Lys Lys Glu Lys Gly Leu Glu Cys
 95 100 105
 Ser Asp Trp Lys Pro Glu Ala Gly Leu Ser Pro Pro Arg Lys Lys
 110 115 120
 Arg Glu Asp Pro Lys Thr Arg Arg Asp Ser Val Asp Ser Lys Ser
 125 130 135
 Ser Ala Ser Ser Ser Pro Lys Arg Pro Ser Val Glu Arg Ser Asn
 140 145 150
 Ser Ser Lys Ser Lys Ala Glu Ser Pro Lys Thr Pro Ser Ser Pro
 155 160 165
 Leu Thr Pro Thr Phe Ala Ser Ser Met Cys Leu Leu Ala Pro Cys
 170 175 180
 Tyr Leu Thr Gly Asp Ser Val Arg Asp Lys Cys Val Glu Met Leu
 185 190 195
 Ser Ala Ala Leu Lys Ala Asp Asp Asp Tyr Lys Asp Tyr Gly Val
 200 205 210
 Asn Cys Asp Lys Met Ala Ser Glu Ile Glu Asp His Ile Tyr Gln
 215 220 225
 Glu Leu Lys Ser Thr Asp Met Lys Tyr Arg Asn Arg Val Arg Ser
 230 235 240
 Arg Ile Ser Asn Leu Lys Asp Pro Arg Asn Pro Gly Leu Arg Arg
 245 250 255
 Asn Val Leu Ser Gly Ala Ile Ser Ala Gly Leu Ile Ala Lys Met
 260 265 270
 Thr Ala Glu Glu Met Ala Ser Asp Glu Leu Arg Glu Leu Arg Asn
 275 280 285
 Ala Met Thr Gln Glu Ala Ile Arg Glu His Gln Met Ala Lys Thr
 290 295 300
 Gly Gly Thr Thr Thr Asp Leu Phe Gln Cys Ser Lys Cys Lys Lys
 305 310 315
 Lys Asn Cys Thr Tyr Asn Gln Val Gln Thr Arg Ser Ala Asp Glu
 320 325 330
 Pro Met Thr Thr Phe Val Leu Cys Asn Glu Cys Gly Asn Arg Trp

Lys Phe Cys 335

340

345

<210> 4
 <211> 201
 <212> PRT
 <213> Homo sapiens

<220> -
 <223> 933353

<400> 4
 Met Ala Ala Thr Ala Leu Leu Glu Ala Gly Leu Ala Arg Val Leu
 1 5 10 15
 Phe Tyr Pro Thr Leu Leu Tyr Thr Leu Phe Arg Gly Lys Val Pro
 20 25 30
 Gly Arg Ala His Arg Asp Trp Tyr His Arg Ile Asp Pro Thr Val
 35 40 45
 Leu Leu Gly Ala Leu Pro Leu Arg Ser Leu Thr Arg Gln Leu Val
 50 55 60
 Gln Asp Glu Asn Val Arg Gly Val Ile Thr Met Asn Glu Glu Tyr
 65 70 75
 Glu Thr Arg Phe Leu Cys Asn Ser Ser Gln Glu Trp Lys Arg Leu
 80 85 90
 Gly Val Glu Gln Leu Arg Leu Ser Thr Val Asp Met Thr Gly Ile
 95 100 105
 Pro Thr Leu Asp Asn Leu Gln Lys Gly Val Gln Phe Ala Leu Lys
 110 115 120
 Tyr Gln Ser Leu Gly Gln Cys Val Tyr Val His Cys Lys Ala Gly
 125 130 135
 Arg Ser Arg Ser Ala Thr Met Val Ala Ala Tyr Leu Ile Gln Val
 140 145 150
 His Lys Trp Ser Pro Glu Glu Ala Val Arg Ala Ile Ala Lys Ile
 155 160 165
 Arg Ser Tyr Ile His Ile Arg Pro Gly Gln Leu Asp Val Leu Lys
 170 175 180
 Glu Phe His Lys Gln Ile Thr Ala Arg Ala Thr Lys Asp Gly Thr
 185 190 195
 Phe Val Ile Ser Lys Thr
 200

<210> 5
 <211> 431
 <212> PRT
 <213> Homo sapiens

<220>
 <221> unsure
 <222> 218
 <223> unknown, or other

<220> -
 <223> 1568361

<400> 5
 Met Ser Ser Val Glu Glu Asp Asp Tyr Asp Thr Leu Thr Asp Ile
 1 5 10 15
 Asp Ser Asp Lys Asn Val Ile Arg Thr Lys Gln Tyr Leu Tyr Val
 20 25 30

Ala	Asp	Leu	Ala	Arg	Lys	Asp	Lys	Arg	Val	Leu	Arg	Lys	Lys	Tyr	
				35					40					45	
Gln	Ile	Tyr	Phe	Trp	Asn	Ile	Ala	Thr	Ile	Ala	Val	Phe	Tyr	Ala	
				50					55					60	
Leu	Pro	Val	Val	Gln	Leu	Val	Ile	Thr	Tyr	Gln	Thr	Val	Val	Asn	
				65					70					75	
Val	Thr	Gly	Asn	Gln	Asp	Ile	Cys	Tyr	Tyr	Asn	Phe	Leu	Cys	Ala	
				80					85					90	
His	Pro	Leu	Gly	Asn	Leu	Ser	Ala	Phe	Asn	Asn	Ile	Leu	Ser	Asn	
				95					100					105	
Leu	Gly	Tyr	Ile	Leu	Leu	Gly	Leu	Leu	Phe	Leu	Leu	Ile	Ile	Leu	
				110					115					120	
Gln	Arg	Glu	Ile	Asn	His	Asn	Arg	Ala	Leu	Leu	Arg	Asn	Asp	Leu	
				125					130					135	
Cys	Ala	Leu	Glu	Cys	Gly	Ile	Pro	Lys	His	Phe	Gly	Leu	Phe	Tyr	
				140					145					150	
Ala	Met	Gly	Thr	Ala	Leu	Met	Met	Glu	Gly	Leu	Leu	Ser	Ala	Cys	
				155					160					165	
Tyr	His	Val	Cys	Pro	Asn	Tyr	Thr	Asn	Phe	Gln	Phe	Asp	Thr	Ser	
				170					175					180	
Phe	Met	Tyr	Met	Ile	Ala	Gly	Leu	Cys	Met	Leu	Lys	Leu	Tyr	Gln	
				185					190					195	
Lys	Arg	His	Pro	Asp	Ile	Asn	Ala	Ser	Ala	Tyr	Ser	Ala	Tyr	Ala	
				200					205					210	
Cys	Leu	Ala	Ile	Val	Ile	Phe	Xaa	Ser	Val	Leu	Gly	Val	Val	Phe	
				215					220					225	
Gly	Lys	Gly	Asn	Thr	Ala	Phe	Trp	Ile	Val	Phe	Ser	Ile	Ile	His	
				230					235					240	
Ile	Ile	Ala	Thr	Leu	Leu	Ser	Thr	Gln	Leu	Tyr	Tyr	Met	Gly		
				245					250					255	
Arg	Trp	Lys	Leu	Asp	Ser	Gly	Ile	Phe	Arg	Arg	Ile	Leu	His	Val	
				260					265					270	
Leu	Tyr	Thr	Asp	Cys	Ile	Arg	Gln	Cys	Ser	Gly	Pro	Leu	Tyr	Val	
				275					280					285	
Asp	Arg	Met	Val	Leu	Leu	Val	Met	Gly	Asn	Val	Ile	Asn	Trp	Ser	
				290					295					300	
Leu	Ala	Ala	Tyr	Gly	Leu	Ile	Met	Arg	Pro	Asn	Asp	Phe	Ala	Ser	
				305					310					315	
Tyr	Leu	Leu	Ala	Ile	Gly	Ile	Cys	Asn	Leu	Leu	Tyr	Phe	Ala		
				320					325					330	
Phe	Tyr	Ile	Ile	Met	Lys	Leu	Arg	Ser	Gly	Glu	Arg	Ile	Lys	Leu	
				335					340					345	
Ile	Pro	Leu	Leu	Cys	Ile	Val	Cys	Thr	Ser	Val	Val	Trp	Gly	Phe	
				350					355					360	
Ala	Leu	Phe	Phe	Phe	Phe	Gln	Gly	Leu	Ser	Thr	Trp	Gln	Lys	Thr	
				365					370					375	
Pro	Ala	Glu	Ser	Arg	Glu	His	Asn	Arg	Asp	Cys	Ile	Leu	Leu	Asp	
				380					385					390	
Phe	Phe	Asp	Asp	His	Asp	Ile	Trp	His	Phe	Leu	Ser	Ser	Ile	Ala	
				395					400					405	
Met	Phe	Gly	Ser	Phe	Leu	Val	Leu	Leu	Thr	Leu	Asp	Asp	Asp	Leu	
				410					415					420	
Asp	Thr	Val	Gln	Arg	Asp	Lys	Ile	Tyr	Val	Phe					
				425					430						

<210> 6
 <211> 419
 <212> PRT
 <213> Homo sapiens

<220> -
 <223> 1600438

<400> 6

Met	Asn	Lys	His	Gln	Lys	Pro	Val	Leu	Thr	Gly	Gln	Arg	Phe	Lys
1				5					10					15
Thr	Arg	Lys	Arg	Asp	Glu	Lys	Glu	Lys	Phe	Glu	Pro	Thr	Val	Phe
				20					25					30
Arg	Asp	Thr	Leu	Val	Gln	Gly	Leu	Asn	Glu	Ala	Gly	Asp	Asp	Leu
				35					40					45
Glu	Ala	Val	Ala	Lys	Phe	Leu	Asp	Ser	Thr	Gly	Ser	Arg	Leu	Asp
				50					55					60
Tyr	Arg	Arg	Tyr	Ala	Asp	Thr	Leu	Phe	Asp	Ile	Leu	Val	Ala	Gly
				65					70					75
Ser	Met	Leu	Ala	Pro	Gly	Gly	Thr	Arg	Ile	Asp	Asp	Gly	Asp	Lys
				80					85					90
Thr	Lys	Met	Thr	Asn	His	Cys	Val	Phe	Ser	Ala	Asn	Glu	Asp	His
				95					100					105
Glu	Thr	Ile	Arg	Asn	Tyr	Ala	Gln	Val	Phe	Asn	Lys	Leu	Ile	Arg
				110					115					120
Arg	Tyr	Lys	Tyr	Leu	Glu	Lys	Ala	Phe	Glu	Asp	Glu	Met	Lys	Lys
				125					130					135
Leu	Leu	Leu	Phe	Leu	Lys	Ala	Phe	Ser	Glu	Thr	Glu	Gln	Thr	Lys
				140					145					150
Leu	Ala	Met	Leu	Ser	Gly	Ile	Leu	Leu	Gly	Asn	Gly	Thr	Leu	Pro
				155					160					165
Ala	Thr	Ile	Leu	Thr	Ser	Leu	Phe	Thr	Asp	Ser	Leu	Val	Lys	Glu
				170					175					180
Gly	Ile	Ala	Ala	Ser	Phe	Ala	Val	Lys	Leu	Phe	Lys	Ala	Trp	Met
				185					190					195
Ala	Glu	Lys	Asp	Ala	Asn	Ser	Val	Thr	Ser	Ser	Leu	Arg	Lys	Ala
				200					205					210
Asn	Leu	Asp	Lys	Arg	Leu	Leu	Glu	Leu	Phe	Pro	Val	Asn	Arg	Gln
				215					220					225
Ser	Val	Asp	His	Phe	Ala	Lys	Tyr	Phe	Thr	Asp	Ala	Gly	Leu	Lys
				230					235					240
Glu	Leu	Ser	Asp	Phe	Leu	Arg	Val	Gln	Gln	Ser	Leu	Gly	Thr	Arg
				245					250					255
Lys	Glu	Leu	Gln	Lys	Glu	Leu	Gln	Glu	Arg	Leu	Ser	Gln	Glu	Cys
				260					265					270
Pro	Ile	Lys	Glu	Val	Val	Leu	Tyr	Val	Lys	Glu	Glu	Met	Lys	Arg
				275					280					285
Asn	Asp	Leu	Pro	Glu	Thr	Ala	Val	Ile	Gly	Leu	Leu	Trp	Thr	Cys
				290					295					300
Ile	Met	Asn	Ala	Val	Glu	Trp	Asn	Lys	Lys	Glu	Glu	Leu	Val	Ala
				305					310					315
Glu	Gln	Ala	Leu	Lys	His	Leu	Lys	Gln	Tyr	Ala	Pro	Leu	Leu	Ala
				320					325					330
Val	Phe	Ser	Ser	Gln	Gly	Gln	Ser	Glu	Leu	Ile	Leu	Leu	Gln	Lys
				335					340					345
Val	Gln	Glu	Tyr	Cys	Tyr	Asp	Asn	Ile	His	Phe	Met	Lys	Ala	Phe
				350					355					360
Gln	Lys	Ile	Val	Val	Leu	Phe	Tyr	Lys	Ala	Asp	Val	Leu	Ser	Glu
				365					370					375
Glu	Ala	Ile	Leu	Lys	Trp	Tyr	Lys	Glu	Ala	His	Val	Ala	Lys	Gly
				380					385					390
Lys	Ser	Val	Phe	Leu	Asp	Gln	Met	Lys	Lys	Phe	Val	Glu	Trp	Leu
				395					400					405
Gln	Asn	Ala	Glu	Glu	Glu	Ser	Glu	Ser	Glu	Gly	Glu	Glu	Asn	
				410					415					

<210> 7

<211> 152

<212> PRT

<213> Homo sapiens

<220> -
 <223> 1605720

<400> 7
 Met Glu Ala Val Leu Asn Glu Leu Val Ser Val Glu Asp Leu Leu
 5 10 15
 Lys Phe Glu Lys Lys Phe Gln Ser Glu Lys Ala Ala Gly Ser Val
 20 25 30
 Ser Lys Ser Thr Gln Phe Glu Tyr Ala Trp Cys Leu Val Arg Ser
 35 40 45
 Lys Tyr Asn Asp Asp Ile Arg Lys Gly Ile Val Leu Leu Glu Glu
 50 55 60
 Leu Leu Pro Lys Gly Ser Lys Glu Glu Gln Arg Asp Tyr Val Phe
 65 70 75
 Tyr Leu Ala Val Gly Asn Tyr Arg Leu Lys Glu Tyr Glu Lys Ala
 80 85 90
 Leu Lys Tyr Val Arg Gly Leu Leu Gln Thr Glu Pro Gln Asn Asn
 95 100 105
 Gln Ala Lys Glu Leu Glu Arg Leu Ile Asp Lys Ala Met Lys Lys
 110 115 120
 Asp Gly Leu Val Gly Met Ala Ile Val Gly Gly Met Ala Leu Gly
 125 130 135
 Val Ala Gly Leu Ala Gly Leu Ile Gly Leu Ala Val Ser Lys Ser
 140 145 150
 Lys Phe

<210> 8
 <211> 811
 <212> PRT
 <213> Homo sapiens

<220> -
 <223> 2515476CD1

<400> 8
 Met Pro Leu Ser Ser Pro Asn Ala Ala Ala Thr Ala Ser Asp Met
 1 5 10 15
 Asp Lys Asn Ser Gly Ser Asn Ser Ser Ser Ala Ser Ser Gly Ser
 20 25 30
 Ser Lys Gly Gln Gln Pro Pro Arg Ser Ala Ser Ala Gly Pro Ala
 35 40 45
 Gly Glu Ser Lys Pro Lys Ser Asp Gly Lys Asn Ser Ser Gly Ser
 50 55 60
 Lys Arg Tyr Asn Arg Lys Arg Glu Leu Ser Tyr Pro Lys Asn Glu
 65 70 75
 Ser Phe Asn Asn Gln Ser Arg Arg Ser Ser Ser Gln Lys Ser Lys
 80 85 90
 Thr Phe Asn Lys Met Pro Pro Gln Arg Gly Gly Gly Ser Ser Lys
 95 100 105
 Leu Phe Ser Ser Ser Phe Asn Gly Gly Arg Arg Asp Glu Val Ala
 110 115 120
 Glu Ala Gln Arg Ala Glu Phe Ser Pro Ala Gln Phe Ser Gly Pro
 125 130 135
 Lys Lys Ile Asn Leu Asn His Leu Leu Asn Phe Thr Phe Glu Pro
 140 145 150
 Arg Gly Gln Thr Gly His Phe Glu Gly Ser Gly His Gly Ser Trp
 155 160 165
 Gly Lys Arg Asn Lys Trp Gly His Lys Pro Phe Asn Lys Glu Leu
 170 175 180
 Phe Leu Gln Ala Asn Cys Gln Phe Val Val Ser Glu Asp Gln Asp
 185 190 195
 Tyr Thr Ala His Phe Ala Asp Pro Asp Thr Leu Val Asn Trp Asp

Phe Val Glu Gln	Val Arg Ile Cys Ser	His Glu Val Pro Ser	Cys
200	205	210	
215	220	225	
Pro Ile Cys Leu	Tyr Pro Pro Thr Ala	Ala Lys Ile Thr Arg	Cys
230	235	240	
Gly His Ile Phe	Cys Trp Ala Cys Ile	Leu His Tyr Leu Ser	Leu
245	250	255	
Ser Glu Lys Thr	Trp Ser Lys Cys Pro	Ile Cys Tyr Ser Ser	Val
260	265	270	
His Lys Lys Asp	Leu Lys Ser Val Val	Ala Thr Glu Ser His	Gln
275	280	285	
Tyr Val Val Gly	Asp Thr Ile Thr Met	Gln Leu Met Lys Arg	Glu
290	295	300	
Lys Gly Val Leu	Val Ala Leu Pro Lys	Ser Lys Trp Met Asn	Val
305	310	315	
Asp His Pro Ile	His Leu Gly Asp Glu	Gln His Ser Gln Tyr	Ser
320	325	330	
Lys Leu Leu Leu	Ala Ser Lys Glu Gln	Val Leu His Arg Val	Val
335	340	345	
Leu Glu Glu Lys	Val Ala Leu Glu Gln	Gln Leu Ala Glu Glu	Lys
350	355	360	
His Thr Pro Glu	Ser Cys Phe Ile Glu	Ala Ala Ile Gln Glu	Leu
365	370	375	
Lys Thr Arg Glu	Glu Ala Leu Ser Gly	Leu Ala Gly Ser Arg	Arg
380	385	390	
Glu Val Thr Gly	Val Val Ala Ala Leu	Glu Gln Leu Val Leu	Met
395	400	405	
Ala Pro Leu Ala	Lys Glu Ser Val Phe	Gln Pro Arg Lys Gly	Val
410	415	420	
Leu Glu Tyr Leu	Ser Ala Phe Asp Glu	Glu Thr Thr Glu Val	Cys
425	430	435	
Ser Leu Asp Thr	Pro Ser Arg Pro Leu	Ala Leu Pro Leu Val	Glu
440	445	450	
Glu Glu Glu Ala	Val Ser Glu Pro Glu	Pro Glu Gly Leu Pro	Glu
455	460	465	
Ala Cys Asp Asp	Leu Glu Leu Ala Asp	Asp Asn Leu Lys Glu	Gly
470	475	480	
Thr Ile Cys Thr	Glu Ser Ser Gln Gln	Glu Pro Ile Thr Lys	Ser
485	490	495	
Gly Phe Thr Arg	Leu Ser Ser Ser Pro	Cys Tyr Tyr Phe Tyr	Gln
500	505	510	
Ala Glu Asp Gly	Gln His Met Phe Leu	His Pro Val Asn Val	Arg
515	520	525	
Cys Leu Val Arg	Glu Tyr Gly Ser Leu	Glu Arg Ser Pro Glu	Lys
530	535	540	
Ile Ser Ala Thr	Val Val Glu Ile Ala	Gly Tyr Ser Met Ser	Glu
545	550	555	
Asp Val Arg Gln	Arg His Arg Tyr Leu	Ser His Leu Pro Leu	Thr
560	565	570	
Cys Glu Phe Ser	Ile Cys Glu Leu Ala	Leu Gln Pro Pro Val	Val
575	580	585	
Ser Lys Glu Thr	Leu Glu Met Phe Ser	Asp Asp Ile Glu Lys	Arg
590	595	600	
Lys Arg Gln Arg	Gln Lys Lys Ala Arg	Glu Glu Arg Arg Arg	Glu
605	610	615	
Arg Arg Ile Glu	Ile Glu Glu Asn Lys	Lys Gln Gly Lys Tyr	Pro
620	625	630	
Glu Val His Ile	Pro Leu Glu Asn Leu	Gln Gln Phe Pro Ala	Phe
635	640	645	
Asn Ser Tyr Thr	Cys Ser Ser Asp Ser	Ala Leu Gly Pro Thr	Ser
650	655	660	
Thr Glu Gly His	Gly Ala Leu Ser Ile	Ser Pro Leu Ser Arg	Ser
665	670	675	
Pro Gly Ser His	Ala Asp Phe Leu Leu	Thr Pro Leu Ser Pro	Thr
680	685	690	

Ala	Ser	Gln	Gly	Ser	Pro	Ser	Phe	Cys	Val	Gly	Ser	Leu	Glu	Glu	
				695					700					705	
Asp	Ser	Pro	Phe	Pro	Ser	Phe	Ala	Gln	Met	Leu	Arg	Val	Gly	Lys	
				710					715					720	
Ala	Lys	Ala	Asp	Val	Trp	Pro	Lys	Thr	Ala	Pro	Lys	Lys	Asp	Glu	
				725					730					735	
Asn	Ser	Leu	Val	Pro	Pro	Ala	Pro	Val	Asp	Ser	Asp	Gly	Glu	Ser	
				740					745					750	
Asp	Asn	Ser	Asp	Arg	Val	Pro	Val	Pro	Ser	Phe	Gln	Asn	Ser	Phe	
				755					760					765	
Ser	Gln	Ala	Ile	Glu	Ala	Ala	Phe	Met	Lys	Leu	Asp	Thr	Pro	Ala	
				770					775					780	
Thr	Ser	Asp	Pro	Leu	Ser	Glu	Glu	Lys	Gly	Gly	Lys	Lys	Arg	Lys	
				785					790					795	
Lys	Gln	Lys	Gln	Lys	Leu	Leu	Phe	Ser	Thr	Ser	Val	Val	His	Thr	
				800					805					810	

Lys

<210> 9
 <211> 352
 <212> PRT
 <213> Homo sapiens

<220> -
 <223> 2754573

<400> 9

Met	His	Val	Val	Ala	Pro	Ala	Ser	Leu	Arg	Leu	Gly	Thr	Gly	Thr	
1				5					10					15	
Asn	Leu	Pro	Pro	Ser	Pro	Thr	Cys	Leu	Thr	Lys	Leu	Ala	Leu	Pro	
				20					25					30	
Pro	Ala	Ala	Glu	Pro	Ser	Leu	Leu	Ala	Met	Ser	Gln	Ser	Arg	His	
				35					40					45	
Arg	Ala	Glu	Ala	Pro	Pro	Leu	Glu	Arg	Glu	Asp	Ser	Gly	Thr	Phe	
				50					55					60	
Ser	Leu	Gly	Lys	Met	Ile	Thr	Ala	Lys	Pro	Gly	Lys	Thr	Pro	Ile	
				65					70					75	
Gln	Val	Leu	His	Glu	Tyr	Gly	Met	Lys	Thr	Lys	Asn	Ile	Pro	Val	
				80					85					90	
Tyr	Glu	Cys	Glu	Arg	Ser	Asp	Val	Gln	Ile	His	Val	Pro	Thr	Phe	
				95					100					105	
Thr	Phe	Arg	Val	Thr	Val	Gly	Asp	Ile	Thr	Cys	Thr	Gly	Glu	Gly	
				110					115					120	
Thr	Ser	Lys	Lys	Leu	Ala	Lys	His	Arg	Ala	Ala	Glu	Ala	Ala	Ile	
				125					130					135	
Asn	Ile	Leu	Lys	Ala	Asn	Ala	Ser	Ile	Cys	Phe	Ala	Val	Pro	Asp	
				140					145					150	
Pro	Leu	Met	Pro	Asp	Pro	Ser	Lys	Gln	Pro	Lys	Asn	Gln	Leu	Asn	
				155					160					165	
Pro	Ile	Gly	Ser	Leu	Gln	Glu	Leu	Ala	Ile	His	His	Gly	Trp	Arg	
				170					175					180	
Leu	Pro	Glu	Tyr	Thr	Leu	Ser	Gln	Glu	Gly	Gly	Pro	Ala	His	Lys	
				185					190					195	
Arg	Glu	Tyr	Thr	Thr	Ile	Cys	Arg	Leu	Glu	Ser	Phe	Met	Glu	Thr	
				200					205					210	
Gly	Lys	Gly	Ala	Ser	Lys	Lys	Gln	Ala	Lys	Arg	Asn	Ala	Ala	Glu	
				215					220					225	
Lys	Phe	Leu	Ala	Lys	Phe	Ser	Asn	Ile	Ser	Pro	Glu	Asn	His	Ile	
				230					235					240	
Ser	Leu	Thr	Asn	Val	Val	Gly	His	Ser	Leu	Gly	Cys	Thr	Trp	His	
				245					250					255	
Ser	Leu	Arg	Asn	Ser	Pro	Gly	Glu	Lys	Ile	Asn	Leu	Leu	Lys	Arg	

	260		265		270
Ser Leu Leu Ser	Ile Pro Asn Thr Asp	Tyr Ile Gln Leu Leu Ser			
	275		280		285
Glu Ile Ala Lys	Glu Gln Gly Phe Asn	Ile Thr Tyr Leu Asp Ile			
	290		295		300
Asp Glu Leu Ser	Ala Asn Gly Gln Tyr	Gln Cys Leu Ala Glu Leu			
	305		310		315
Ser Thr Ser Pro	Ile Thr Val Cys His	Gly Ser Gly Ile Ser Cys			
	320		325		330
Gly Asn Ala Gln	Ser Asp Ala Ala His	Asn Ala Leu Gln Tyr Leu			
	335		340		345
Lys Ile Ile Ala	Glu Arg Lys				
	350				

<210> 10
 <211> 432
 <212> PRT
 <213> Homo sapiens

<220> -
 <223> 2926777

<400> 10

Met Ile Ser Ala Ala	Gln Leu Leu Asp	Glu Leu Met Gly Arg Asp
1	5	10
Arg Asn Leu Ala Pro	Asp Glu Lys Arg Thr	Asn Val Arg Trp Asp
	20	25
His Glu Ser Val Cys	Lys Tyr Tyr Leu Cys	Gly Phe Cys Pro Ala
	35	40
Glu Leu Phe Thr Asn	Thr Arg Ser Asp Leu	Gly Pro Cys Glu Lys
	50	55
Ile His Asp Glu Asn	Leu Arg Lys Gln Tyr	Glu Lys Ser Ser Arg
	65	70
Phe Met Lys Val Gly	Tyr Glu Arg Asp Phe	Leu Arg Tyr Leu Gln
	80	85
Ser Leu Leu Ala Glu	Val Glu Arg Arg Ile	Arg Arg Gly His Ala
	95	100
Arg Leu Ala Leu Ser	Gln Asn Gln Gln Ser	Ser Gly Ala Ala Gly
	110	115
Pro Thr Gly Lys Asn	Glu Glu Lys Ile Gln	Val Leu Thr Asp Lys
	125	130
Ile Asp Val Leu Leu	Gln Gln Ile Glu Glu	Leu Gly Ser Glu Gly
	140	145
Lys Val Glu Glu Ala	Gln Gly Met Met Lys	Leu Val Glu Gln Leu
	155	160
Lys Glu Glu Arg Glu	Leu Leu Arg Ser Thr	Thr Ser Thr Ile Glu
	170	175
Ser Phe Ala Ala Gln	Glu Lys Gln Met Glu	Val Cys Glu Val Cys
	185	190
Gly Ala Phe Leu Ile	Val Gly Asp Ala Gln	Ser Arg Val Asp Asp
	200	205
His Leu Met Gly Lys	Gln His Met Gly Tyr	Ala Lys Ile Lys Ala
	215	220
Thr Val Glu Glu Leu	Lys Glu Lys Leu Arg	Lys Arg Thr Glu Glu
	230	235
Pro Asp Arg Asp Glu	Arg Leu Lys Lys Glu	Lys Gln Glu Arg Glu
	245	250
Glu Arg Glu Lys Glu	Arg Glu Arg Glu Arg	Glu Glu Arg Glu Arg
	260	265
Lys Arg Arg Arg Glu	Glu Glu Glu Arg Glu	Lys Glu Arg Ala Arg
	275	280
Asp Arg Glu Arg Arg	Lys Arg Ser Arg Ser	Arg Ser Arg His Ser

	290		295		300
Ser Arg Thr Ser	Asp Arg Arg Cys Ser	Arg Ser Arg Asp His	Lys		
	305		310		315
Arg Ser Arg Ser	Arg Glu Arg Arg Arg	Thr Arg Ser Arg Asp	Arg		
	320		325		330
Arg Arg Ser Arg	Ser His Asp Arg Ser	Glu Arg Lys His Arg	Ser		
	335		340		345
Arg Ser Arg Asp	Arg Arg Arg Ser Lys	Ser Arg Asp Arg Lys	Ser		
	350		355		360
Tyr Lys His Arg	Ser Lys Ser Arg Asp	Arg Glu Gln Asp Arg	Lys		
	365		370		375
Ser Lys Glu Lys	Glu Lys Arg Gly Ser	Asp Asp Lys Lys Ser	Ser		
	380		385		390
Val Lys Ser Gly	Ser Arg Glu Lys Gln	Ser Glu Asp Thr Asn	Thr		
	395		400		405
Glu Ser Lys Glu	Ser Asp Thr Lys Asn	Glu Val Asn Gly Thr	Ser		
	410		415		420
Glu Asp Ile Lys	Ser Glu Gly Asp Thr	Gln Ser Asn			
	425		430		

<210> 11
 <211> 1750
 <212> DNA
 <213> Homo sapiens

<220> -
 <223> 009476

<400> 11

gcccgcattggt	aactcaggcg	ccggggcgca	tgctctagct	gctgggttttc	cacgctgggt	60
ttagctccccg	gcgtctgcaa	aatgaagatt	gaggaggtga	agagcactac	gaagacgcag	120
cgcctcgcct	cccacagcca	cgtgaaagg	ctggggctgg	acgagagcgg	cttgcccaag	180
caggcggcct	cagggttgt	gggcccagg	aacgcgcgag	aggcatgtgg	cgctcatagta	240
gaattaatcg	aaagcaagaa	aatggctgga	agagctgtct	tggtggcagg	acctcctgga	300
actggcaaga	cagctctggc	tctggctatt	gctcaggagc	tggttagtaa	ggtccccctc	360
tgcccaatgg	tggggagtga	agtttactca	actgagatca	agaagacaga	ggtgctgatg	420
gagaacttcc	gcaggggccat	tgggctgcga	ataaaggaga	ccaagggaagt	ttatgaaggt	480
gaagtcacag	agctaactcc	gtgtgagaca	gagaatccca	tgggaggata	tgccaaaacc	540
attagccatg	tgatcatagg	actcaaaaca	gccaaaggaa	ccaaacagtt	gaaactggac	600
cccagcattt	ttgaaagtgt	gcagaaagag	cgagttagaag	ctggagatgt	gatttacatt	660
gaagccaaca	gtggggccgt	gaagaggcag	ggcagggtgt	atacctatgc	cacagaattc	720
gaccttgaag	ctgaagagta	tgtccctttg	ccaaaagggg	atgtgcacaa	aaagaaagaa	780
atcatccaag	atgtgacctt	gcattgactt	gatgtggcta	atgcgcggcc	ccagggggga	840
caagatatcc	tgtccatgat	gggccagcta	atgaagccaa	agaagacaga	aatcacagac	900
aaacttcgag	gggagattaa	taagggtggt	aacaagtaca	tcgaccagg	cattgctgag	960
ctgggtccccg	gtgtgctgtt	tgttgatgag	gtccacatgc	tggacattga	gtgcttcacc	1020
tacctgcacc	gcgccttgg	gtcttctatc	gtcccatcgc	tcattctttgc	atccaaccga	1080
ggcaactgtg	tcattcagagg	cactgaggac	atcacatccc	ctcacggcat	ccctcttgac	1140
cttctggacc	gagtgatgat	aatccggacc	atgctgtata	ctccacagga	aatgaaacag	1200
atcattaaaa	tccgtgcccc	gacggaagga	atcaacatca	gtgaggaggc	actgaaccac	1260
ctggggggaga	ttggcaccac	gaccacactg	aggtactcag	tgcagctgct	gaccccgcc	1320
aacttgcttg	ctaaaatcaa	cggaaggac	agcattgaga	aagagcatgt	cgaagagatc	1380
agtgaacttt	tctatgatgc	caagtctctc	gccaaaatcc	tggtgacca	gcaggataag	1440
tacatgaagt	gagatggctg	aggttttcag	cagtaagaga	ctccccagg	gtgcttgcc	1500
tggttccagc	ctgtgggcgc	ttgcccctgg	gcttggggct	gccgtcccca	ctcaggcgtg	1560
gtctgcagcg	ctgtcagttc	agtgtggaaa	gcatttcttt	ttaaagttatc	gtaactgttc	1620
ctgtgggttg	tttgaaagaa	cccttctcta	cctgggtgtg	tttctataaa	tcttcatagg	1680
ttattttgat	tctctctctc	tctctctctc	agttttttta	aaataaactt	ttcagaacag	1740
aaaaaaaaaa						1750

<210> 12

<211> 691
<212> DNA
<213> Homo sapiens

<220> -
<223> 077180

<400> 12
gggaaaaatgg cgtctggccat gctgggtcttg gtgggtttcgc cgtgggtctgc ggccccgggga 60
gtgtcttcgaa actactggga gcgactgcta cggaagcttc cgcagagccg gccgggcttt 120
cccagtcctc cgtggggacc agcattagca gtacagggcc cagccatgtt tacagagcca 180
gcaaatgata ccagtgggaag taaagagaat tccagccttt tggacagtat cttttggatg 240
gcagctccca aaaatagacc caccattgaa gttaaccggt gtaggagaag aaatccgcag 300
aagcttatta aagttaagaa caacatagac gtttgtcctg aatgtggtca cctgaaacag 360
aaacatgtcc tttgtgccta ctgctatgaa aaggtgtgca aggagactgc agaaatcaga 420
cgacagatag ggaagcaaga agggggccct ttttaaggctc ccaccataga gactgtggtg 480
ctgtacacag gagagacacc gtctgaacaa gatcagggca agaggatcat tgaacgagac 540
agaaagcgac catcctgggt caccacagaat tgacacccaa gatgttaaaa ggataacttc 600
acagtaaatc atttctcctg aaatagagga agattcttta cgttgtttgtg cttgttttta 660
aatcatcagt atagtttaac acattctttc t 691

<210> 13
<211> 1594
<212> DNA
<213> Homo sapiens

<220> -
<223> 118160

<400> 13
ccgccccgc cgggcgtgtg tgtcgtgtgt gtttggggcc cgcgcgggtt gcgcgccttc 60
cgccttcgcg cctcctgccc ccgaggccct actgctgccc ctgtgcccct cgccccgcgcg 120
ggcgtcgcgg gccaacatgg gccaggaaga ggagctgctg aggatcgcca aaaagctgga 180
gaagatggtg gccaggaaga acacggaagg ggccctggac cttctgaaga agctgcacag 240
ctgccagatg tccatccagc tactacagac aaccaggatt ggagttgctg ttaatggggg 300
ccgcaagcac tgcctagaca aggaggtggt gtccctggcc aaagtcctta tcaaaaactg 360
gaagcggctg ctgactccc ctggaccccc aaaaggagaa aaaggagagg aaagagaaaa 420
ggcaaaagaag aaggaaaaaag ggcttgagtg ttcagactgg aagccagaag caggcctttc 480
tccaccaagg aaaaaacgag aagaccccaa aaccaggaga gactctgtgg actccaagtc 540
ttctgcctcc tccctcccaa aaagaccatc ggtggaaaga tcaaacagca gcaaatcaaa 600
agcggagagc cccaaaacac ctgacagccc cttgaccccc acgtttgcct cttccatgtg 660
tctcctggcc cctctctatc tcacagggga ctctgtccgg gacaagtgtg tggagatgct 720
gtcagcagcc ctgaaggcgg acgatgatta caaggactat ggagtcaact gtgacaagat 780
ggcatcagaa atcgaagatc atatctacca agagctcaag agcacggaca tgaagtaccg 840
gaaccgcgtg cgcagccgca taagcaacct caaggacccc aggaaccccc gcttgcggcg 900
gaacgtgctc agtggggcca tctccgcagg gcttatagcc aagatgacgg cagaggaaat 960
ggccagtgat gaactgaggg agttgaggaa tgccatgacc caggaggcca tccgtgagca 1020
ccagatggcc aagactggcg gcaccaccac tgacctcttc cagtgcagca aatgcaagaa 1080
gaagaactgc acctataacc aggtgcagac acgcagtgtc gatgagcca tgactacctt 1140
tgtcttatgc aatgaatgtg gcaatcgctg gaagttctgc tgatggaaca gccagccatg 1200
aacaagggtg ggaagaagaa agaggaagcg ctgaattatc tgaactggag aagcaataaa 1260
aattaaagtg aaggaaaata ctgaactctg tctgagtggg atggtatgag ttagaggaag 1320
aattctcttg caaattaata atcggtcatt agaaaacaatt ggttaaatggg ggagcctaata 1380
tgagagaatg tgcctgagaat ttgtattgat gaacctcttt tagaaactgc agagggctgg 1440
gcacgggtggc ttatggctgt aatctgcaaa ctctgggagg ctgaggtggg agaactcgctt 1500
aacccagaa gtttgagtc agcccaggca acacagcaag accccatctc tataaaaaaga 1560
aaaaataaag aaattgtaga cgcctcgggg acat 1594

<210> 14
<211> 821
<212> DNA
<213> Homo sapiens

<220> -

<223> 933353

<400> 14

```

cgggtcgccg ctgcgccggg cgggatggc ggcaccgcg ctgctggagg cgggctggc 60
gcggtgctc ttctaccga cgtgctcta caccctgtc cgcgggaagg tgccgggtcg 120
ggcgaccgg gactgggtacc accgcatcga cccaccgtg ctgctggggc cgtgcccgtt 180
gcgagcttg acgcgccagc tggtagagga cgagaacgtg cgcgggggtga tcaccatgaa 240
cgaggagtag gagacgaggt tctgtgcaa ctcttcacag gaggtagaaga gactaggagt 300
tgagcagctg cggctcagca cagtagacat gactgggatc cccaccttgg acaacctcca 360
gaagggagtc caatttgctc tcaagtacca gtcgctgggc cagtgtgttt acgtgcattg 420
taaggctggg cgctccagga gtgccactat ggtggcagca tactgtattc aggtgcacaa 480
atggagtcca gaggaggctg taagagccat cgccaagatc cggtcataca tccacatcag 540
gcttgccag ctggatgttc ttaaagagtt ccacaagcag attactgcac gggcaacaaa 600
ggatgggact ttgttcattt caaagacatg atgtatgggg attagaaaga actcaagaca 660
ctcctgcttg atacagaaca aaaagagctt aacaggacca acagggttta agcccagact 720
tgacgtaaca gaaatgtgcc aataggtaat aggtaatatt tctttctctg acttgttttg 780
ttttcttgaa ataacactgt tgtgtggcta gaaaaaaaaa a 821

```

<210> 15

<211> 2028

<212> DNA

<213> Homo sapiens

<220>

<221> unsure

<222> 2, 4, 6, 7, 15, 18, 19, 41, 59, 70, 74, 95, 97, 119, 127,

<221> unsure

<222> 131, 152, 158, 889

<223> a or g or c or t, unknown, or other

<220> -

<223> 1568361

<400> 15

```

gnanannacc ctgcnngnng ccattcacgg accctgccc nacagccgtc accctcgant 60
cctggctgan tctnttcttg gcagttcccc ttatnanggt tacaactatg gctcctttna 120
caatgtntct ntatctaccg atgggtctgg tncagcncct ggcaactggg accctcttta 180
cggttaccag ggccgctcct ttgaacctgt aggtactcgg ccccgagtgg actccatgag 240
ctctgtggag gaggatgact acgacacatt gaccgacatc gattccgaca agaattgtcat 300
tgccaccaag caataacctct atgtggctga cctggcacgg aaggacaagc gtgttctgcg 360
gaaaaagtac cagatctact tctggaacat tgccaccatt gctgtcttct atgccccttc 420
tgtgtgtcag ctggtgatca cctaccagac ggtgggtgaat gtcacaggga atcaggacat 480
ctgctactac aacttctctt gcgcccacc actgggcaat ctacagcct tcaacaacat 540
ctcagcaaac ctgggggtaca tctgtctggg gctgcttttc ctgctcatca tctgcaaacg 600
ggagatcaac cacaaccggg cctgtctgcg caatgacctc tgtgccctgg aatgtgggat 660
ccccaaacac ttggggcttt tctacgccat gggcacagcc ctgatgatgg aggggctgct 720
cagtgcttgc tatcatgtgt gccccaacta taccaatttc cagtttgaca catcgttcat 780
gtacatgatc gccggactct gcatgctgaa gctctaccag aagcggcacc cggacatcaa 840
cgccagcgcc tacagtgcct acgctgcct ggccattgtc atcttcttnt ctgtgctggg 900
cgtggtcttt ggcaaaaggga acacggcggt ctggatcgtc ttctccatca ttcacatcat 960
gcgccacctg ctctcagca cgcagctcta ttacatggg cggtggaac tgactcggg 1020
gatcttccgc cgcatectcc acgtgctcta cacagactgc atccggcagt gcagcgggcc 1080
gctctacgtg gaccgcatgg tgctgctggg catgggcaac gtcacaaact ggtcgtggc 1140
tgccataggg ctatcatgc gccccaatga ttctgcttcc tactgttg ccattggcat 1200
ctgcaacctg ctcttttact tcgcttcta catcatcatg aagctccgga gtggggagag 1260
gatcaagctc atccccctgc tctgcacctc ttgcacctcc gtggtctggg gcttcgcgtc 1320
cttcttcttc ttccaggac tcagcacctg gcagaaaacc cctgcagagt cgaggagca 1380
caaccgggac tgcatectcc tcgacttctt tgacgaccac gacatctggc acttctctc 1440
ctccatcgcc atgttcgggt ccttctctgg gtgtctgaca ctggatgacg acctggatac 1500
tgtgcagcgg gacaagatct atgtcttcta gcaggagctg ggcccttcgc ttcacctcaa 1560
ggggccctga gctcttttgt gtcatagacc ggtcactctg tctgtctgtg gggatgagtc 1620
ccagcacgcg tgccagcac tggatggcag caggacagcc aggtctagct taggcttggc 1680
ctgggacagc catgggggtgg catggaacct tgcagctgcc ctctgccag gacagggcct 1740

```



```

gctccccctgg aacccccaga tgttggccaa attgctgctt tcttctcagt gttggggcct 1800
tccatggggcc cctgtccctt ggctctccat ttgtcccttt gcaagaggaa ggatggaagg 1860
gacaccctcc ccatttcatg ccttgcattt tgcccgctct cctccccaca atgccccagc 1920
ctgggaccta aggctctt ttctcccat actccactc cagggcctag tctggggcct 1980
gaatctctgt cctgtatcag ggcgccagtt ctctttgggc tgtccctg 2028

```

<210> 16
 <211> 1830
 <212> DNA
 <213> Homo sapiens

<220> -
 <223> 1600438

```

<400> 16
gtcaatgtgt ctgtcccttca ctccctccatt gtctgccgcc actgctgctg ctgctgctgc 60
tgccgctgct gctgcacgaa tcgtcgcagc ccccagcctt gcgcgtcgtc gctacctct 120
cggacagaaa ttttatgaat aagcatcaga agccagtgtt aacaggccag cggttcaaaa 180
ctcggaaaaa ggatgaaaaa gagaaattcg aaccacagct cttcagggat acacttgtcc 240
aggggcttaa tgaggctggt gatgaccttg aagctgtagc caaatttctg gactctacag 300
gctcaagatt agattatcgt cgctatgcag acacactctt cgatatcctg gtggctggca 360
gtatgcttgc ccctggagga acgcgcatag atgatggtga caagaccaag atgaccaacc 420
actgtgtgtt ttcagcaaat gaagatcatg aaaccatccg aaactatgct cagggtcttca 480
ataaactcat caggagatat aagtatttgg agaaggcatt tgaagatgaa atgaaaaagc 540
ttctcctctt ccttaaaagcc ttttccgaaa cagagcagac aaagtggcg atgctgtcgg 600
ggattctgct gggcaatggc tccctgcccg ccaccatcct caccagtctc ttcaccgaca 660
gcttagtcaa agaaggcatt aggcctcat ttgctgtcaa gcttttcaaa gcatggatgg 720
cagaaaaaga tgccaactct gttacctcgt ctttgagaaa agccaactta gacaagaggc 780
tgcttgaact ctttccagtt aacagacaga gtgtggatca ttttgctaaa tacttctactg 840
acgcaggctt taaggagctt tccgacttcc tccgagtcca gcagtccctg ggcaccagga 900
aggaactgca gaaggagctc caggagcgtc tttctcagga atgcccgatc aaggagggtg 960
tgctttatgt caaagaagaa atgaagagga atgatcttcc agaaacagca gtgattggtc 1020
ttctgtggac atgtataatg aacgctgttg agtggaaaca gaaggaaaga cttgttgcag 1080
agcaggctct gaagcacctg aagcaatatg ctcccctgct ggccgtgttc agctcccaag 1140
gccagtcaga gctgatcctc ctccagaagg ttcaggaaata ctgctacgac aacatccatt 1200
tcatgaaaagc ctttcagaag attgtggttc tcttttataa agctgatgtt ctgagcgaag 1260
aagcaatact gaaatggtat aaggaagcac atgttgctaa aggcaaaagt gtttttcttg 1320
accagatgaa gaaatttgtt gagtgggtac aaaatgcaga agaagaatcc gaatcggaa 1380
gtgaggaaaa ttaaatggct caacaagcac aatacctagg ttaccacaca ccactttttg 1440
attgggaatg ctgaaccatt tgagaagaga aacttggctt ctgttttcgc aaaggaaaaa 1500
aaaaatagga taggcctccc ttgtgcagag ggagaaaatg ttttgttttt gttttgtttt 1560
taaattggagc cctgagggcat cagctattat acttgggact ctacctctca ctactatat 1620
gctaacttaa agccattcaa caaggagtc aatagatctg aaattaaata ctcaacagac 1680
tctcctcttt ttagctgtat ttttcaggta ctgtgtggtg accgccccac tgggtgtctat 1740
tacaggccac tttggtagtt gtgtatctgc tcatgtatgt gatttgacaa accagttttt 1800
taaaataaat ggcttttttaa gaaaaaaaaa 1830

```

<210> 17
 <211> 566
 <212> DNA
 <213> Homo sapiens

<220> -
 <223> 1605720

```

<400> 17
gttctccgcc cctgccactg ggccatggag actgtggcac agtagactgt agtgtgaggc 60
tcgcgggggc agtgcccatg gagggcgtgc tgaacgagct ggtgtctgtg gaggacctgc 120
tgaagtgtga aaagaaattt cagtctgaga aggcagcagg ctcggtgtcc aagagcacgc 180
agtgtgagta cgcctgggtc ctggtgcgga gcaagtacaa tgatgacatc cgtaaaggca 240
tcgtgctgct cgaggagctg ctgccc aaag ggagcaagga ggaacagcgg gattacgtct 300
tctacctggc cgtgggggaa taccggctca aggaatacga gaaggcctta aagtacgtcc 360
gcgggttgc gcagacagag ccccgagaaca accaggccaa ggaactggag cggctcattg 420

```

acaaggccat gaagaaagat ggactcgtgg gcatggccat cgtgggaggg atggccctgg 480
 gtgtggcggg acttgccgga ctcatcgga cttgtgtgtc caagtccaaa ttctgaagga 540
 gacgcgggag cccacggaga acgctc 566

<210> 18
 <211> 3520
 <212> DNA
 <213> Homo sapiens

<220> -
 <223> 2515476

<400> 18
 gagaagccaa ggaaggaaac agggaaaaat gtcgccatga aggccgagaa ccgctgccgc 60
 cgccgacccc cgccggccct gaacgccatg agcctgggtc cccgccgcgc ccgctccgct 120
 tcgactgccc tcgcccgcga ggcccccggt gatgccgtg agctccccc acgccgcgcg 180
 caccgcctcc gacatggaca agaacagcgg ctccaacagc tcctccgctt cttcgggcag 240
 cagcaaaagg caacagccgc cccgctccgc ctggcgggg ccagccggcg agtctaaacc 300
 caagagcgat ggaagaact ccagtggatc caagcgttat aatcgcaaac gtgaactttc 360
 ctaccccaaa aatgaaagt ttaacaacca gtcccgctgc tcagttcac agaaaagcaa 420
 gacttttaac aagatgcctc ctcaaagggg cggcggcagc agcaaaactt ttagctcttc 480
 ttttaattgtt ggaagacgag atgaggtagc agaggctcaa cgggcagagt ttagccctgc 540
 ccagttctct ggtcctaaga agatcaacct gaaccacttg ttgaatttca cttttgaacc 600
 ccgtggccag acgggtcact ttgaaggcag tggacatggt agctggggaa agaggaacaa 660
 gtggggacat aagcctttta acaaggaaact ctttttacag gccaaactgc aatttgggt 720
 gcttgaagac caagactaca cagctcattt tctgtatcct gatacattag ttaactggga 780
 ctttgtgga caagtgcgca ttgttagcca tgaagtgcc tcttgcccaa tatgctcta 840
 tccacctact gcagccaaga taacccgttg tggacacatc ttctgctggg catgcatcct 900
 gcactatctt tcaactgagt agaagacgtg gagtaaatgt cccatctgtt acagtctctg 960
 gcataaagaag gatctcaaga gtgttgttgc cacagagtca catcagtatg ttgttgggtg 1020
 taccattacg atgcagctga tgaagaggga gaaaggggtg ttgggtggctt tgcccaaate 1080
 caaatggatg aatgtagacc atcccattca tgaaggagt gctgcacccg gtatgtctgg 1140
 caagtgtctg ctggcctcta aggagcaggt gctgcacccg gtatgtctgg aggagaaagt 1200
 agcactagag cagcagctgg cagaggagaa gcacactccc gactcctgct ttattgaggc 1260
 agctatccag gagctcaaga ctcggaaga ggctctgtcg ggattggccg gaagcagaag 1320
 ggaggtcact ggtgttgttg ctgctctgga acaactggtg ctgatggctc ccttggcgaa 1380
 ggagtctgtt tttaaccca ggaagggtgt gctggagtat ctgtctgctc tcgatgaaga 1440
 aaccacggaa gtttgttctc tggacactcc ttctagacct cttgctctcc ctctggtaga 1500
 agaggaggaa gcagtgtctg aaccagagcc tgagggggtg ccagaggcct gtgatgactt 1560
 ggagttagca gatgacaatc ttaagagggg gaccatttgc actgagttca gccagcagga 1620
 accctacac aagtcaggct tcacacgctc cagcagctct ccttgttact acttttacca 1680
 agcggaagat ggacagcata tgttctgca cctgtgaaat gtgcgctgcc tegtgcggga 1740
 gtacggcagc ctggagagga gcccgcagaa gatctcagca actgtggtgg agattgtgg 1800
 ctactccatg tctgaggatg ttcgacagcg tcacagatat ctctctcact tgccactcac 1860
 ctgtgagttc agcatctgtg aactggcttt gcaacctcct gtggtctcta aggaaacctt 1920
 agagatgttc tcagatgaca ttgagaagag gaaacgtcag cgccaaaaga aggtcggga 1980
 ggaacgccgc cgagagcgca ggattgagat agaggagaac aagaaaacag gcaagtacct 2040
 agaagtccac attccctctg agaattctaca gcagtttctt gccttcaatt cttataacct 2100
 ctctctgat tctgttttg gtcccaccag caccgagggc catggggccc tctccatttc 2160
 tcctctcagc agaagtccag gttcccatgc agactttctg ctgacccctc tgtcaccac 2220
 tgccagtcag ggcagtcctt cattctgcgt tgggagttct gaagaagact ctcccttccc 2280
 ttcccttggc cagatgctga gggttgaaa agcaaaaagca gatgtgtggc ccaaaactgc 2340
 tccaaaagaa gatgagaaca gcttagttcc tcttgcctt gtggacagcg acggggagag 2400
 tgataattca gaccgtgttc ctgtgcccag ttttcaaaat tcttccagcc aagctattga 2460
 agcagccttc atgaaactgg acacaccagc tacttcagat cccctctctg aagagaaagg 2520
 aggaaagaaa agaaaaaac agaaacagaa gctcctgttc agcacctcag tcgtccacac 2580
 caagtgcac tactggccca ggctaccttc tccacttggt ttttgtttt gttttttt 2640
 ccccatgct tttgtttggc tgctgtaatt ttaagtatt tgagtttgaa cagattagct 2700
 ctgggggggag ggggtttcca caatgtgagg ggaaccaag aaaattttta atacagtga 2760
 ttttccagct tctgtcttt acacaaaaat aaagtattga cacaagagat ctcttctgc 2820
 caaggttttt agttcattgc cagtttagtc ttttgacct atgtgtaatt aatttttctc 2880
 aacccaaagt aagattgagt ccccttttag atgcattaga gcagttccaa ccagaatggc 2940
 acacactgct ctgctgtacc atcagtcag ggcttctgg actcagtaca cctctcagtt 3000
 tgtcttttaa aaaacagctg aatctttact acctatttag ttctccttgc taaagaaaca 3060

```

gggggtgggaa taaaatggat ttaggacacc cagtttgaat tgcagttttt tttttctga 3120
cacatggcca ggctgtgggt ccagcttaat ggagtaggct gtccttggca cttgcatgtg 3180
tgaaaggagg gttttgcctc ttcttgagca tggcttgagt tggtaaggaa agctgtaact 3240
cacgaagccc tgagacctgc taccocctaa atcgagcttg ttttcagtga ctggcttgag 3300
tcataggagg aggagtctgg tacagctgca ggagagcagg gccatctgaa gcggtagcat 3360
tgccaccatc tccctctcat cttagagcagt ttctttatgc cttgggttga gctgaatttg 3420
atgtgaattc ttttgctgct taataaagtg acctctaggt gcattagaat gcgaaggcaa 3480
atagttgcaa taaatcacct gcacaagcaa aaaaaaaaaa 3520

```

```

<210> 19
<211> 1904
<212> DNA
<213> Homo sapiens

<220>
<221> unsure
<222> 32, 33
<223> a or g or c or t, unknown, or other

```

```

<220> -
<223> 2754573

```

```

<400> 19
cgggggaggtc ggaggagggtg gcggcgctgg anntcctccc ggggaccagc gaccggggag 60
catgcacgtc gtgcaccag cttcactgag gtttgaaca ggaactaatc tccctccatc 120
tccaacttgc ttgacaaagc tgcctcttcc tccagccgct gagccgtccc ttctcgccat 180
gtcccagagc aggcaaccgc ccgaggcccc gccgctggag cgcgaggaca gtgggacctt 240
cagttttgggg aagatgataa cagctaagcc agggaaaaca ccgattcagg tattacacga 300
atcggcatg aagaccaaga acatcccagt ttatgaatgt gaaagatctg atgtgcaaat 360
acacgtgccc actttcacct tcagagtaac cgttgggtgac ataacctgca cagggtgaagg 420
tacaagtaag aagctggcga aacatagagc tgcagaggct gccataaaca ttttgaaagc 480
caatgcaagt atttgctttg cagttcctga ccccttaatg cctgacctt ccaagcaacc 540
aaagaaccag cttaatccta ttggttcatt acaggaattg gctattcact atggctggag 600
acttccctgaa tatacccttt cccaggaggg aggcctgct cataagagag aatatactac 660
aatttgcagg cttagagtcatt ttatggaaac tggaaagggg gcatcaaaaa agcaagccaa 720
aaggaatgct gctgagaaat ttcttgccaa atttagtaat atttctccag agaaccacat 780
ttctttaaca aatgtagtag gacattcttt aggatgtact tggcattcct tgaggaattc 840
tcttggtgaa aagatcaact tactgaaaag aagcctcctt agtattccaa atacagatta 900
catccagctg cttagtgaaa ttgccaagga acaagggttt aatataacat atttggatat 960
agatgaactg agcgccaatg gacaatatca atgtcttgct gaactgtcca ccagccccat 1020
cacagtctgt catggctccg gtatctctctg tggcatgca caaagtgatg cagctcaca 1080
tgctttgcag tatttaaaaga taatagcaga aagaaagtaa atctggagca acttaaaaaa 1140
tctttcagta gcacataaaa agttcccttc tggcccttcc ccaagtaaaa cttttaccgt 1200
agtgtttatg tcttgtttct aaatctcttc atagattcca tcaacactcc agatttaatt 1260
atctctcat agttgttatt aagctctttt taatggcttc aactttgtat cagtatactg 1320
tatttataaa ctttgtagca caagagagag tgtagcacc attttacagt gccatgcaca 1380
tcagagaaag aaactgcatg tttgttggtg atgatgaaat aaaaatgcta gcgacagtct 1440
ttcttactgg tgcttaagct cttctttgca caaagcttta taaagggaat tcaaagggaag 1500
ccctttagaa ttagagtcct gagggacagc actaacaggc ctttattaag tatgattgat 1560
tgttaaattt cagggaacat gattggctct cgtgtattt gaattcatgt aacaaagaac 1620
tgttacgatg ggattctgct cattttatta aaaagctact gacttgactg tcatcctgtt 1680
cttgtagacc attgtgaata agattttaat gttgataatt ctgttattta catatctcta 1740
atttactttg aaattcaaaag gtgaaaataa aaaatgatgg cctaagtaaa atttacaac 1800
atacatgtaa tgggttactt ccttactttc tctaagcgta tgcaaccttg tatttttctt 1860
tgacataacc atacacaggt gaattactag tttaaaaagc attt 1904

```

```

<210> 20
<211> 1621
<212> DNA
<213> Homo sapiens

```

```

<220> -
<223> 2926777

```

<400> 20

at tt t g g g t t t	t t a a a g g g a a	g g c g t c c g c g	c g g c g g c c a t	t t t g t c t t g t	c g g c t c c t g t	60
g t g t a g g a g g	g a t t t c g g c c	t g a g a g c g g g	c c g a g g a g a t	t g g c g a c g g t	g t c g c c c g t g	120
c t t t t c g t t g	g c g g g t g c c t	g g g c t g g t g g	g a a c a g c c g c	c c g a a g g a a g	c a c c a t g a t t	180
t c g g c c g c g c	a g t t g t t g g a	t g a g t t a a t g	g g c c g g g a c c	g a a a c c t a g c	c c c g g a c g a g	240
a a g c g c a c a a	a c g t g c g g t g	g g a c c a c g a g	a g c g t t t g t a	a a t a t t a t c t	c t g t g g t t t t	300
t g t c c t g c g g	a a t t g t t c a c	a a a t a c a c g t	t c t g a t c t t g	g t c c g t g t g a	a a a a a t t c a t	360
g a t g a a a a t c	t a c g a a a a c a	g t a t g a g a a g	a g c t c t c g t t	t c a t g a a a g t	t g g c t a t g a g	420
a g a g a t t t t t	t g c g a t a c t t	a c a g a g c t t a	c t t g c a g a a g	t a g a a c g t a g	g a t c a g a c g a	480
g g c c a t g c t c	g t t t g g c a t t	a t c t c a a a a c	c a g c a g t c t t	c t g g g g c c g c	t g g c c c a a c a	540
g g c a a a a a t g	a a g a a a a a a t	t c a g g t t c t a	a c a g a c a a a a	t t g a t g t a c t	t c t g c a a c a g	600
a t t g a a g a a t	t a g g g t c t g a	a g g a a a a g t a	g a a g a a g c c c	a g g g g a t g a t	g a a a t t a g t t	660
g a g c a a t t a a	a a g a a g a g a g	a g a a c t g c t a	a g g t c c a c a a	c g t c g a c a a t	t g a a a g c t t t	720
g c t g c a c a a g	a a a a a c a a a t	g g a a g t t t g t	g a a g t a t g t g	g a g c c t t t t t	a a t a g t a g g a	780
g a t g c c c a g t	c c c g g g t a g a	t g a c c a t t t g	a t g g g a a a a c	a a c a c a t g g g	c t a t g c c a a a	840
a t t a a a g c t a	c t g t a g a a g a	a t t a a a a g a a	a a g t t a a g g a	a a a g a a c c g a	a g a a c c t g a t	900
c g t g a t g a g c	g t c t a a a a a a	g g a g a a g c a a	g a a a g a g a a g	a a a g a g a a a a	a g a a c g g g a g	960
a g a g a a a g g g	a a g a a a g a g a	a a g g a a a a g a	c g a a g g g a a g	a g g a a g a a a g	a g a a a a a g a a	1020
a g g g c t c g t g	a c a g a g a a a g	a a g a a a g a g a	a g t c g t t c a c	g a a g t a g a c a	c t c a a g c c g a	1080
a c a t c a g a c a	g a a g a t g c a g	c a g g t c t c g g	g a c c a c a a a a	g g t c a c g a a g	t a g a g a a a g a	1140
a g g c g g a c c a	g a a g t a g a g a	t c g a c g a a g a	a g c a g a a g c c	a t g a t c g a t c	a g a a a g a a a a	1200
c a c a g a t c t c	g a a g t c g g g a	t c g a a g a a g a	t c a a a a a g c c	g g g a t c g a a a	g t c a t a t a a g	1260
c a c a g g a g c a	a a a g t c g g g a	c a g a g a a c a a	g a t a g a a a a t	c c a a g g a g a a	a g a a a a g a g g	1320
g g a t c t g a t g	a t a a a a a a a g	t a g t g t g a a g	t c c g g t a g t c	g a g a a a a g c a	g a g t g a a g a c	1380
a c a a a c a c t g	a a t c g a a g g a	a a g t g a t a c t	a a g a a t g a g g	t c a a t g g g a c	c a g t g a a g a c	1440
a t t a a a t c t g	a a g g t g a c a c	t c a g t c c a a t	t a a a a c t g a t	c t g a t a a g a c	c t c a g a t c a g	1500
a c a g a g g a c t	a c t g t t c g a a	g a t t t t t g g a	a g a a t a c t g a	g a a c g g c a t a	a a g t g a a g a t	1560
c g a c a t t t a a	a a a a t g a g g t	g a a a g a a a g c	t a t a g t g g c a	t a g a a a a a g t	a t a a a g c t c a	1620
g						1621



INTERNATIONAL APPLICATION PUBLISHED UNDER THE PATENT COOPERATION TREATY (PCT)

(51) International Patent Classification ⁶ : C12N 15/12, C07K 14/47, C12N 15/70, 1/21, A61K 38/17, C07K 16/18, C12Q 1/68	A3	(11) International Publication Number: WO 99/15658 (43) International Publication Date: 1 April 1999 (01.04.99)
(21) International Application Number: PCT/US98/19839 (22) International Filing Date: 22 September 1998 (22.09.98) (30) Priority Data: 08/933,750 23 September 1997 (23.09.97) US (63) Related by Continuation (CON) or Continuation-in-Part (CIP) to Earlier Application US 08/933,750 (CIP) Filed on 23 September 1997 (23.09.97) (71) Applicant (for all designated States except US): INCYTE PHARMACEUTICALS, INC. [US/US]; 3174 Porter Drive, Palo Alto, CA 94304 (US). (72) Inventors; and (75) Inventors/Applicants (for US only): LAL, Preeti [IN/US]; 2382 Lass Drive, Santa Clara, CA 95054 (US). HILLMAN, Jennifer, L. [US/US]; 230 Monroe Drive #12, Mountain View, CA 94040 (US). BANDMAN, Olga [US/US]; 366 Anna Avenue, Mountain View, CA 94043 (US). SHAH, Purvi [IN/US]; 1608 Queen Charlotte Drive #5, Sunnyvale, CA 94087 (US). AU-YOUNG, Janice [US/US]; 1419 Kains		Avenue, Berkeley, CA 94702 (US). YUE, Henry [US/US]; 826 Lois Avenue, Sunnyvale, CA 94087 (US). GUEGLER, Karl, J. [CH/US]; 1048 Oakland Avenue, Menlo Park, CA 94025 (US). (74) Agents: BILLINGS, Lucy, J. et al.; Incyte Pharmaceuticals, Inc., 3174 Porter Drive, Palo Alto, CA 94304 (US). (81) Designated States: AL, AM, AT, AU, AZ, BA, BB, BG, BR, BY, CA, CH, CN, CU, CZ, DE, DK, EE, ES, FI, GB, GE, GH, GM, HR, HU, ID, IL, IS, JP, KE, KG, KP, KR, KZ, LC, LK, LR, LS, LT, LU, LV, MD, MG, MK, MN, MW, MX, NO, NZ, PL, PT, RO, RU, SD, SE, SG, SI, SK, SL, TJ, TM, TR, TT, UA, UG, US, UZ, VN, YU, ZW, ARIPO patent (GH, GM, KE, LS, MW, SD, SZ, UG, ZW), Eurasian patent (AM, AZ, BY, KG, KZ, MD, RU, TJ, TM), European patent (AT, BE, CH, CY, DE, DK, ES, FI, FR, GB, GR, IE, IT, LU, MC, NL, PT, SE), OAPI patent (BF, BJ, CF, CG, CI, CM, GA, GN, GW, ML, MR, NE, SN, TD, TG). Published <i>With international search report.</i> <i>Before the expiration of the time limit for amending the claims and to be republished in the event of the receipt of amendments.</i> (88) Date of publication of the international search report: 10 June 1999 (10.06.99)
(54) Title: HUMAN REGULATORY MOLECULES (57) Abstract <p>The invention provides human regulatory molecules and polynucleotides (collectively designated HRM) which identify and encode them. The invention also provides expression vectors, host cells, agonists, antibodies and antagonists. The invention further provides methods for diagnosing, preventing, and treating disorders associated with expression of human regulatory molecules.</p>		

FOR THE PURPOSES OF INFORMATION ONLY

Codes used to identify States party to the PCT on the front pages of pamphlets publishing international applications under the PCT.

AL	Albania	ES	Spain	LS	Lesotho	SI	Slovenia
AM	Armenia	FI	Finland	LT	Lithuania	SK	Slovakia
AT	Austria	FR	France	LU	Luxembourg	SN	Senegal
AU	Australia	GA	Gabon	LV	Latvia	SZ	Swaziland
AZ	Azerbaijan	GB	United Kingdom	MC	Monaco	TD	Chad
BA	Bosnia and Herzegovina	GE	Georgia	MD	Republic of Moldova	TG	Togo
BB	Barbados	GH	Ghana	MG	Madagascar	TJ	Tajikistan
BE	Belgium	GN	Guinea	MK	The former Yugoslav Republic of Macedonia	TM	Turkmenistan
BF	Burkina Faso	GR	Greece			TR	Turkey
BG	Bulgaria	HU	Hungary	ML	Mali	TT	Trinidad and Tobago
BJ	Benin	IE	Ireland	MN	Mongolia	UA	Ukraine
BR	Brazil	IL	Israel	MR	Mauritania	UG	Uganda
BY	Belarus	IS	Iceland	MW	Malawi	US	United States of America
CA	Canada	IT	Italy	MX	Mexico	UZ	Uzbekistan
CF	Central African Republic	JP	Japan	NE	Niger	VN	Viet Nam
CG	Congo	KE	Kenya	NL	Netherlands	YU	Yugoslavia
CH	Switzerland	KG	Kyrgyzstan	NO	Norway	ZW	Zimbabwe
CI	Côte d'Ivoire	KP	Democratic People's Republic of Korea	NZ	New Zealand		
CM	Cameroon			PL	Poland		
CN	China	KR	Republic of Korea	PT	Portugal		
CU	Cuba	KZ	Kazakhstan	RO	Romania		
CZ	Czech Republic	LC	Saint Lucia	RU	Russian Federation		
DE	Germany	LI	Liechtenstein	SD	Sudan		
DK	Denmark	LK	Sri Lanka	SE	Sweden		
EE	Estonia	LR	Liberia	SG	Singapore		

INTERNATIONAL SEARCH REPORT

International Application No.

PC1/US 98/19839

A. CLASSIFICATION OF SUBJECT MATTER

IPC 6 C12N15/12 C07K14/47 C12N15/70 C12N1/21 A61K38/17
C07K16/18 C12Q1/68

According to International Patent Classification (IPC) or to both national classification and IPC

B. FIELDS SEARCHED

Minimum documentation searched (classification system followed by classification symbols)

IPC 6 C12N C07K A61K C12Q

Documentation searched other than minimum documentation to the extent that such documents are included in the fields searched

Electronic data base consulted during the international search (name of data base and, where practicable, search terms used)

C. DOCUMENTS CONSIDERED TO BE RELEVANT

Category *	Citation of document, with indication, where appropriate, of the relevant passages	Relevant to claim No.
X	MASATO KANEMAKI ET AL.: "Molecular cloning of a rat 49-kDa TBP-interacting protein (TIP49) that is highly homologous to the bacterial RuvB" BIOCHEMICAL AND BIOPHYSICAL RESEARCH COMMUNICATIONS, vol. 235, no. 1, 9 June 1997, pages 64-68, XP002088338 ORLANDO, FL US see abstract see page 65, left-hand column, paragraph 6 - page 68, left-hand column, paragraph 1; figure 1 -----	1-14,17, 18,21-24

☐ Further documents are listed in the continuation of box C.

☐ Patent family members are listed in annex.

* Special categories of cited documents:

- "A" document defining the general state of the art which is not considered to be of particular relevance
- "E" earlier document but published on or after the international filing date
- "L" document which may throw doubts on priority claim(s) or which is cited to establish the publication date of another citation or other special reason (as specified)
- "O" document referring to an oral disclosure, use, exhibition or other means
- "P" document published prior to the international filing date but later than the priority date claimed

- "T" later document published after the international filing date or priority date and not in conflict with the application but cited to understand the principle or theory underlying the invention
- "X" document of particular relevance; the claimed invention cannot be considered novel or cannot be considered to involve an inventive step when the document is taken alone
- "Y" document of particular relevance; the claimed invention cannot be considered to involve an inventive step when the document is combined with one or more other such documents, such combination being obvious to a person skilled in the art.
- "&" document member of the same patent family

Date of the actual completion of the international search

17 December 1998

Date of mailing of the international search report

03.05.99

Name and mailing address of the ISA

European Patent Office, P.B. 5818 Patentstein 2
NL - 2280 HV Rijswijk
Tel (+31-70) 340-3040, Tx. 31 651 apo nl,
Fax (+31-70) 340-3016

Authorized officer

MONTERO LOPEZ B.

INTERNATIONAL SEARCH REPORT

national application No.

PCT/US 98/ 19839

Box I Observations where certain claims were found unsearchable (Continuation of item 1 of first sheet)

This International Search Report has not been established in respect of certain claims under Article 17(2)(a) for the following reasons:

1. ☒ Claims Nos.:
because they relate to subject matter not required to be searched by this Authority, namely:

Please see Further Information sheet enclosed.
2. ☒ Claims Nos.:
because they relate to parts of the International Application that do not comply with the prescribed requirements to such an extent that no meaningful International Search can be carried out, specifically:

Please see Further Information sheet enclosed.
3. ☐ Claims Nos.:
because they are dependent claims and are not drafted in accordance with the second and third sentences of Rule 6.4(a).

Box II Observations where unity of invention is lacking (Continuation of item 2 of first sheet)

This International Searching Authority found multiple inventions in this international application, as follows:

Please see additional sheet.

1. ☐ As all required additional search fees were timely paid by the applicant, this International Search Report covers all searchable claims.
2. ☐ As all searchable claims could be searched without effort justifying an additional fee, this Authority did not invite payment of any additional fee.
3. ☐ As only some of the required additional search fees were timely paid by the applicant, this International Search Report covers only those claims for which fees were paid, specifically claims Nos.:
4. ☒ No required additional search fees were timely paid by the applicant. Consequently, this International Search Report is restricted to the invention first mentioned in the claims; it is covered by claims Nos.:

Claims 1-24 partially.

Remark on Protest

- ☐ The additional search fees were accompanied by the applicant's protest.
- ☐ No protest accompanied the payment of additional search fees.

FURTHER INFORMATION CONTINUED FROM PCT/ISA/ 210

1. Claims: Partially 1-24

Human regulatory molecule comprising SEQ ID NO:1, polynucleotide comprising SEQ ID NO:11 encoding the polypeptide, polynucleotides hybridizing therewith or complementary thereto; microarray containing a fragment of a polynucleotide of SEQ ID NO:11; vector and host cell comprising the polynucleotides and use thereof to produce the polypeptide; agonists and antagonists of the polypeptides; pharmaceutical composition comprising the human regulatory protein and antibody binding the same; use thereof in stimulating cell proliferation and preventing cancer; use of the polynucleotides an microarray for detecting expression levels of polynucleotides encoding a human regulatory molecule.

2. Claims: Partially 1-24

Human regulatory molecule comprising SEQ ID NO:2, polynucleotide comprising SEQ ID NO:12 encoding the polypeptide, polynucleotides hybridizing therewith or complementary thereto; microarray containing a fragment of a polynucleotide of SEQ ID NO:12; vector and host cell comprising the polynucleotides and use thereof to produce the polypeptide; agonists and antagonists of the polypeptides; pharmaceutical composition comprising the human regulatory protein and antibody binding the same; use thereof in stimulating cell proliferation and preventing cancer; use of the polynucleotides an microarray for detecting expression levels of polynucleotides encoding a human regulatory molecule.

3. Claims: Partially 1-24

Human regulatory molecule comprising SEQ ID NO:3, polynucleotide comprising SEQ ID NO:13 encoding the polypeptide, polynucleotides hybridizing therewith or complementary thereto; microarray containing a fragment of a polynucleotide of SEQ ID NO:13; vector and host cell comprising the polynucleotides and use thereof to produce the polypeptide; agonists and antagonists of the polypeptides; pharmaceutical composition comprising the human regulatory protein and antibody binding the same; use thereof in stimulating cell proliferation and preventing cancer; use of the polynucleotides an microarray for detecting expression levels of polynucleotides encoding a human regulatory molecule.

4. Claims: Partially 1-24

Human regulatory molecule comprising SEQ ID NO:4, polynucleotide comprising SEQ ID NO:14 encoding the

FURTHER INFORMATION CONTINUED FROM PCT/ISA/ 210

polypeptide, polynucleotides hybridizing therewith or complementary thereto; microarray containing a fragment of a polynucleotide of SEQ ID NO:14; vector and host cell comprising the polynucleotides and use thereof to produce the polypeptide; agonists and antagonists of the polypeptides; pharmaceutical composition comprising the human regulatory protein and antibody binding the same; use thereof in stimulating cell proliferation and preventing cancer; use of the polynucleotides an microarray for detecting expression levels of polynucleotides encoding a human regulatory molecule.

5. Claims: Partially 1-24

Human regulatory molecule comprising SEQ ID NO:5, polynucleotide comprising SEQ ID NO:15 encoding the polypeptide, polynucleotides hybridizing therewith or complementary thereto; microarray containing a fragment of a polynucleotide of SEQ ID NO:15; vector and host cell comprising the polynucleotides and use thereof to produce the polypeptide; agonists and antagonists of the polypeptides; pharmaceutical composition comprising the human regulatory protein and antibody binding the same; use thereof in stimulating cell proliferation and preventing cancer; use of the polynucleotides an microarray for detecting expression levels of polynucleotides encoding a human regulatory molecule.

6. Claims: Partially 1-24

Human regulatory molecule comprising SEQ ID NO:6, polynucleotide comprising SEQ ID NO:16 encoding the polypeptide, polynucleotides hybridizing therewith or complementary thereto; microarray containing a fragment of a polynucleotide of SEQ ID NO:16; vector and host cell comprising the polynucleotides and use thereof to produce the polypeptide; agonists and antagonists of the polypeptides; pharmaceutical composition comprising the human regulatory protein and antibody binding the same; use thereof in stimulating cell proliferation and preventing cancer; use of the polynucleotides an microarray for detecting expression levels of polynucleotides encoding a human regulatory molecule.

7. Claims: Partially 1-24

Human regulatory molecule comprising SEQ ID NO:7, polynucleotide comprising SEQ ID NO:17 encoding the polypeptide, polynucleotides hybridizing therewith or complementary thereto; microarray containing a fragment of a polynucleotide of SEQ ID NO:17; vector and host cell comprising the polynucleotides and use thereof to produce

FURTHER INFORMATION CONTINUED FROM PCT/ISA/ 210

the polypeptide; agonists and antagonists of the polypeptide; pharmaceutical composition comprising the human regulatory protein and antibody binding the same; use thereof in stimulating cell proliferation and preventing cancer; use of the polynucleotides an microarray for detecting expression levels of polynucleotides encoding a human regulatory molecule.

8. Claims: Partially 1-24

Human regulatory molecule comprising SEQ ID NO:8, polynucleotide comprising SEQ ID NO:18 encoding the polypeptide, polynucleotides hybridizing therewith or complementary thereto; microarray containing a fragment of a polynucleotide of SEQ ID NO:18; vector and host cell comprising the polynucleotides and use thereof to produce the polypeptide; agonists and antagonists of the polypeptides; pharmaceutical composition comprising the human regulatory protein and antibody binding the same; use thereof in stimulating cell proliferation and preventing cancer; use of the polynucleotides an microarray for detecting expression levels of polynucleotides encoding a human regulatory molecule.

9. Claims: Partially 1-24

Human regulatory molecule comprising SEQ ID NO:9, polynucleotide comprising SEQ ID NO:19 encoding the polypeptide, polynucleotides hybridizing therewith or complementary thereto; microarray containing a fragment of a polynucleotide of SEQ ID NO:19; vector and host cell comprising the polynucleotides and use thereof to produce the polypeptide; agonists and antagonists of the polypeptide; pharmaceutical composition comprising the human regulatory protein and antibody binding the same; use thereof in stimulating cell proliferation and preventing cancer; use of the polynucleotides an microarray for detecting expression levels of polynucleotides encoding a human regulatory molecule.

10. Claims: Partially 1-24

Human regulatory molecule comprising SEQ ID NO:10, polynucleotide comprising SEQ ID NO:20 encoding the polypeptide, polynucleotides hybridizing therewith or complementary thereto; microarray containing a fragment of a polynucleotide of SEQ ID NO:20; vector and host cell comprising the polynucleotides and use thereof to produce the polypeptide; agonists and antagonists of the polypeptide; pharmaceutical composition comprising the human regulatory protein and antibody binding the same; use thereof in stimulating cell proliferation and preventing

FURTHER INFORMATION CONTINUED FROM PCT/ISA/ 210

cancer; use of the polynucleotides an microarray for detecting expression levels of polynucleotides encoding a human regulatory molecule.

FURTHER INFORMATION CONTINUED FROM PCT/ISA/ 206

Although claim 17, as far as concerning an in vivo method, and claim 18 are directed to a method of treatment of the human/animal body, the search has been carried out and based on the alleged effects of the compound/composition.

Further defect(s) under Article 17(2)(a):

Claims Nos.: 15, 16, 19, 20

Claims 15, 16, 19 and 20, referring to agonists and antagonists have not been searched due to the lack of technical description therefore.

